Voting on Multiple Issues: What to Put on the Ballot ?

Alex Gershkov, Benny Moldovanu and Xianwen Shi*

August 29, 2018

Abstract

We study a multi-dimensional collective decision under incomplete information. Agents have Euclidean preferences and vote by simple majority on each issue (dimension), yielding the coordinate-wise median. Judicious rotations of the orthogonal axes – the issues that are voted upon – lead to welfare improvements. If the agents' types are drawn from a distribution with independent marginals then, under weak conditions, voting on the original issues is not optimal. If the marginals are identical (but not necessarily independent), then voting first on the total sum and next on the differences is often welfare superior to voting on the original issues. We also provide various lower bounds on incentive efficiency: in particular, if agents' types are drawn from a log-concave density with I.I.D. marginals, a second-best voting mechanism attains at least 88% of the first-best efficiency. Finally, we generalize our method and some of our insights to preferences derived from distance functions based on inner products.

*A earlier version of the paper was circulated under the title "The Dimensions of Consensus." We wish to thank Yeon-Koo Che, Rahul Deb, Hans-Peter Grüner, Philippe Jehiel, Andreas Kleiner, Robert J. McCann, Konrad Mierendorff, Marcin Peski, Rob Ready, Colin Stewart, Thomas Tröger, Jürgen von Hagen, Cedric Wasser, the co-editor and referees for helpful comments, and to various seminar and conference participants for helpful discussions. Shuangjian Zhang provided excellent research assistance. Gershkov's research is supported by a grant from Israel Science Foundation. Shi is grateful to Social Sciences and Humanities Research Council of Canada for financial support. Gershkov: Department of Economics, Hebrew University of Jerusalem, Israel and School of Economics, University of Surrey, UK, alexg@huji.ac.il; Moldovanu: Department of Economics, University of Bonn, Germany, mold@uni-bonn.de; Shi: Department of Economics, University of Toronto, Canada, xianwen.shi@utoronto.ca.

1 Introduction

In 1974 the U.S. Congress changed its budgeting process: instead of considering appropriations requests that were voted upon one at a time (*bottom-up*) which resulted in a gradually determined total level of spending, the *Congressional Budget and Impoundment Control Act* required voting first on an overall level of spending, before the determination of budgets for individual programs in subsequent votes *(top-down)*. A large literature in the area of public finance (see for example the review articles in Poterba and von Hagen [1999]) has debated the costs and benefits of such procedural changes, with particular attention to the size of the expected budget deficit.[1]

We analyze the problem of redefining (or bundling) the issues brought to vote in a multi-dimensional collective decision problem. Such methods can increase the welfare of the involved decision makers by allowing them to reach a consensus that was not possible on the original issues.

We study a multi-dimensional collective decision taken by simple majority voting: an example is a legislature that needs to decide on individual budgets for public goods such as, say, education and defense. Other examples are decisions on the geographical location of a desirable facility, or decisions on hiring and project adoption that are based on multi-dimensional attributes.

We adopt the standard spatial model of voting widely used in the political science literature (see for example, Chapter 5 in Austen-Smith and Banks [2005]), where voters have preferences characterized by ideal points in each dimension, and by a quadratic loss caused by deviations from the ideal point.[2]

Voters' ideal points are private information, and we study voting by simple majority on each dimension separately. As we shall see below, this focus yields, in combination with a decision over the dimensions that are the subject of voting, an analysis of more generality than immediately apparent.

Voting by simple majority on each dimension yields the coordinate-wise median of the voters' ideal points. This easily follows from Black's [1948] famous theorem because the induced preferences are single-peaked on each one-dimensional issue. In general, this outcome does not coincide with the first-best, the alternative that minimizes the sum of squared distances from the individual ideal points. The first-best is the coordinate-wise average (or mean) of the realized ideal points, and thus first-best welfare is the corresponding variance (with a minus sign).

[1] There was a widespread belief that the new rules would lead to smaller deficits, and the act was passed almost unanimously in both House and Senate.

[2] The main text deals with the two-dimensional case, while the generalization to more than two dimensions is in an Appendix.

The first-best is not implementable: each agent has an incentive to try to move the average closer to his/her ideal point by exaggerating his/her position on one or more issues.[3] Given the tension between first-best on the one hand and implementable outcomes on the other, how well does voting by simple majority perform in terms of welfare? A classical inequality due to Hotelling and Solomons [1932] implies that, for any distribution of preferences, voting by simple majority on any given issues achieves at least 50% of the first-best welfare.

The main insight of the present paper is that a judicious choice of the issues that are actually put to vote (while maintaining voting by simple majority, with its desirable incentive properties) can significantly improve welfare.[4] For example, instead of voting on two separate issues, the legislature could vote on a total budget, and then on a division of that budget between the two issues – just as Congress started to do in 1974. More generally, we model the repackaging and bundling of issues by rotations of the orthogonal axes that define what is put to vote. For example, suppose voters care about two separate main issues, but they actually vote on the budget of two agencies that overlap in their responsibility over these two issues. Rotations correspond then to the shifting of jurisdictions among the two agencies: they change the mix of issues under the control of each agency.

In influential work, Shepsle [1979] argued that the division of a complex decision into several different jurisdictions (*germaneness*), creates stable equilibria that would not be possible in a general, unconstrained collective decision model. His main examples are legislative committees in the U.S. congress. Viewed in light of Shepsle's theory, our goal is to endogenize the choice of jurisdictions in order to improve welfare, an issue that has not received much attention in formal studies.

A basic technical observation is that the mean is rotation equivariant (i.e., the mean after rotation is obtained by rotating the original mean) but the coordinate-wise median is not.[5] As a consequence, a rotation of the axes may decrease the distance between the coordinate-wise mean (first-best) and the coordinate-wise median (outcome of majority voting), thus increasing welfare. The basic cause behind this phenomenon is the non-linearity of the median function, a feature that yields a rather complex analysis.[6] In order to use calculus and probabilistic/statistical techniques,

[3] This observation was first made by Galton [1907], who was also the first to recommend the use of the median as a robust and non-manipulable aggregator of opinions. His insights have been sharpened and much generalized in the literature on *robust estimation*.

[4] The idea of comparing voting rules in terms of their expected welfare goes back to Rae [1969].

[5] See Haldane [1948], or the literature on spatial voting, e.g., Feld and Grofman [1988].

[6] This is true even for common distributions of types, such as the Gamma, Poisson, lognormal, etc. Some of our results are based on insights that go back to conjectures by Ramanujan (see Szegö

we focus here on the limit case where the number of voters is infinite.

Our main results are:

1) If the agents' ideal points in one dimension are independently distributed from the ideal points in the other dimension then, under weak conditions on the distribution of preferences, voting on the original issues is sub-optimal; that is, a re-packaging of the issues brought to vote via rotation (which necessarily creates some correlation among the ideal points) increases welfare. This parallels the non-optimality of separate sales in the multi-product monopoly problem: some form of mixed bundling is always superior to separate sales (see McAfee, McMillan and Whinston [1989]).

2) If the marginals of the distribution of agents' ideal points are identically distributed (not necessarily independently), we provide sufficient conditions under which the 45-degree rotation welfare is superior to no rotation. The conditions are satisfied by common distribution with I.I.D. marginals. We show that, with I.I.D. marginals, the 45-degree rotation is always a critical point, and also provide sufficient conditions for the 45-degree rotation to be welfare maximizing. A key observation for these results is that, under the symmetry of the marginals, the 45-degree rotation entirely eliminates the conflict arising between efficiency and majority voting in one dimension – all remaining conflict is concentrated in the other, orthogonal dimension.

3) We provide various lower bounds on incentive efficiency for large, non-parametric families of distributions of ideal points (such as unimodal distributions, distributions with an increasing hazard rate, etc.). For example, if agents' ideal points are drawn from a log-concave density with I.I.D. marginals, a voting mechanism that involves a 45-degree rotation of the original dimensions attains at least 88% of the first-best efficiency.

4) We extend our method to the more general class of preferences induced by distance functions generated by inner-product norms. In particular, for weighted Euclidean norms, we show that voting on independent issues remains sub-optimal under the same sufficient conditions as for the Euclidean preferences.

It is possible to perform a similar analysis for goals other than efficiency, e.g., define jurisdictions that serve other purposes, such as the self-interest of an agenda setter, or of a coalition of voters. Ferejohn and Krehbiel [1987] focused on controlling budgetary growth rather than efficiency, and they observed that the 1974 budget reform can be represented by a 45-degree rotation of the coordinates on which voting takes place. For that goal, we offer here precise conditions comparing the top-down and bottom-up procedures in terms of the total budget they produce, and we show that the budgeting reform can unambiguously improve welfare while having a **mixed**

[1928]) and Hadamard.

impact on the budget size.

To see how our results may fit practical voting environments, consider a legislative committee that decides on spending on several items. Each committee member has a preferred expenditure for each item. If the items are independent (i.e, the preferred expenditure level on one item is uncorrelated with the preferred one on another item) then it is not optimal to directly vote on the proposed expenditures. Instead, it may be better to vote on the budgets of two agencies that have some overlapping jurisdictions representing a particular mix of the two issues (this is a non-zero rotation in our framework). In another example, if a committee finances regional hospitals, say, that have similar sizes and serve similar purposes, our analysis suggests that it is better to first decide the total budget for these hospitals and then divide it among hospitals. Finally, if a government, say, has to fund an activity for multiple years, it may be better first vote on a multi-year budget and then decide how to allocate the total budget among different years.

1.1 Related Literature

The existence of a Condorcet winner is rare in multi-dimensional models of voting (Kramer [1973]). Kramer [1972] observed, however, that voting in a variety of institutions is often sequential, issue by issue, and he established the existence of a sophisticated voting equilibrium if voters' preferences are continuous, convex and separable. The coordinate-wise median – obtained by simple-majority voting in each dimension – constitutes a basic instance of a *structure induced equilibrium* in the spirit of Shepsle [1979].[7]

Technically, our contribution builds upon and relates to several important and elegant contributions due to Moulin [1980], Border and Jordan [1983], Kim and Rousch [1984], and Peters, van der Stel and Storcken [1992]. In a one-dimensional setting with single-peaked preferences, Moulin considered mechanisms that depend on reported peaks, and characterized the set of dominant strategy incentive compatible (DIC), anonymous and Pareto efficient mechanisms: each mechanism in this class is obtained by choosing the median among the n reported peaks of the real voters and the peaks of a set of $n-1$ "phantom" voters (these are fixed by the mechanism, and do not vary with the reports).[8] Border and Jordan [1983] removed Moulin's assumption whereby mechanisms depend only on peaks, and generalized Moulin's finding

[7] In a multi-dimensional voting model with common interest, aggregate uncertainty, and two truth-motivated candidates, McMurray [2018] shows that, in equilibrium, multiple issues are consistently bundled along the 45-degree line (the major diagonal in his model).

[8] Relaxing Pareto efficiency yields the same characterization, but requires $n+1$ phantoms.

to a multi-dimensional setting with separable and quadratic preferences: each DIC mechanism is decomposable into a collection of one-dimensional DIC mechanisms, each described by the location of the phantom voters in the respective dimension (see also Barbera, Gul and Stacchetti [1993]).[9]

Gershkov, Moldovanu and Shi [2017] analyzed welfare maximization in a one-dimensional setting with cardinal utilities, and derived the ex-ante welfare maximizing placement of phantoms. They also showed how to avoid the phantom interpretation by implementing Moulin's mechanisms via a sequential, binary voting procedure together with a flexible qualified majority schedule.[10] Combining their result with the Border-Jordan decomposition yields the welfare maximizing mechanism for multi-dimensional settings with separable and quadratic preferences. But, the ensuing solution, described by an optimal placement of phantoms in each dimension, is not satisfactory from a practical point of view: it implies that each issue (dimension) in each multi-dimensional problem must be voted upon according to a particular institution that is sensitive to both utilities and distribution of types. Such flexibility may be difficult, if not impossible, to achieve in practice.

Instead, we fix here an ubiquitous institution – voting by simple majority on each issue – but we allow flexibility in the design of the issues that are actually put to vote. Such a limited form of agenda design is common in practice, and, as we shall see, has important welfare consequences.

The simplest multi-dimensional setting is the one with Euclidean preferences: intuitively, the presence of spherically symmetric preferences does not a-priori determine the dimensions of the Border and Jordan decomposition into one-dimensional mechanisms. Indeed, Kim and Rousch [1984] showed that the set of continuous, anonymous and DIC mechanisms can be described by performing the Border-Jordan analysis subsequent to any translation of the origin and any rotation of the orthogonal axes.[11] Peters, van der Stel and Storcken [1992] showed that, for two dimensions with odd number of voters, voting by simple majority in each dimension (after any translation/rotation of the plane) is also Pareto efficient.[12]

[9]Most papers in the literature indeed assume separable preferences. Ahn and Oliveros [2012] is a notable exception: they prove equilibrium existence in combinatorial voting with non-separable preferences, and provide conditions under which the Condorcet winner is implemented in the equilibrium of large elections.

[10]See also Kleiner and Moldovanu [2017] for general sufficient conditions under which sequential, binary voting procedures possess desirable properties.

[11]Since both median and mean are translation equivariant, translations of the origin cannot improve welfare. It is therefore without loss of generality to restrict attention here to rotations.

[12]They show that a mechanism is Pareto efficient if and only if, for any realization of agents' ideal points, its allocation lies in the convex hull of the ideal points. With two or more dimensions, a

Finally, it is also instructive to compare our results to those in the classical papers by Caplin and Nalebuff ([1988], [1991]).[13] These authors did not consider incomplete information and incentive constraints. Instead, motivated by the instability of multi-dimensional voting, they considered the effect of super-majority requirements on the stability of the spatial mean. For a large number of voters and for a log-concave density governing the distribution of types (and also for other, more general forms of concavity), Caplin and Nalebuff showed that, once established as status-quo, the mean cannot be displaced by another alternative if the selection of that alternative requires a super-majority of at least 64% (or $1 - \frac{1}{e}$). In other words, any coalition that prefers an alternative over the mean contains less than 64% of the voters, and is thus not effective.

As mentioned above, for the log-concave case with I.I.D. marginals, our results display a mechanism that is incentive compatible for any (odd) number of voters and that achieves at least 88% of the first-best utility when this number goes to infinity. Thus, issue by issue simple majority voting on appropriately defined dimensions constitutes an intuitive and incentive compatible institutional arrangement that is almost efficient in this case. Moreover, the relative efficiency of this mechanism increases, and tends to 100%, when we increase the number of dimensions of the underlying problem.

Although our setting bears some similarity to multi-dimensional cheap talk, the logic of welfare gains is very different here. In those models, the multiplicity of issues helps because it improves information transmission between the sender(s) and the receiver. In a model with two senders Battaglini [2002] shows that, as long as the two senders' ideal points are linearly independent, full information revelation is possible by carefully choosing dimensions to exploit the conflict between senders. In a one-sender model, Chakraborty and Harbaugh [2007] show that the sender can credibly convey his ranking of different issues to the receiver. In our model rotations address a very different conflict, one between simple majority voting and efficiency.

2 The Model

We consider n (odd) agents who collectively decide about two issues, X and Y, on a convex region $D \subseteq \mathbb{R}^2$. Each agent's ideal position on these two issues is given by a peak $\mathbf{t}_i = (x_i, y_i)$, $i = 1, 2, ..., n$. The peak $\mathbf{t}_i$ is agent i's private information. Each

generalized median with phantoms may lie ouside of the convex hull.

[13] These papers were also the first to use modern concentration inequalities in the Economics literature.

agent i has a utility function of the form

$$-\left|\left|\mathbf{t}_i - \mathbf{v}\right|\right|^2$$

where the point $\mathbf{v} \in D$ denotes the chosen alternative and where $||\cdot||$ is the standard Euclidean (l_2) norm. The peaks $\mathbf{t}_i = (x_i, y_i)$ are independently, identically distributed (I.I.D.) across agents, according to a joint distribution $F(x_i, y_i)$, with density f. Denote by μ_X (μ_Y) the expected value of x_i (y_i). Throughout the paper, we assume that $\mathbb{E}\left|\left|\mathbf{t}_i\right|\right|^2 < \infty$ for all $\mathbf{t}_i$.

A utilitarian planner would choose $\mathbf{v} \in D$ to maximize the average of the agents' ex ante utilities, or equivalently, minimize the expected average squared distance from the voters' peaks:

$$\min_{\mathbf{v}\in D} \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}\left|\left|\mathbf{t}_i - \mathbf{v}\right|\right|^2\right],$$

subject to agents' incentive constraints. Ignoring the agents' incentives, the planner would choose a point $\mathbf{u}$ that minimizes the average of ex post distances:

$$\mathbf{u} \in \arg\min_{\mathbf{v}\in D} \frac{1}{n}\sum_{i=1}^{n}\left|\left|\mathbf{t}_i - \mathbf{v}\right|\right|^2,$$

which we will refer to as the *first-best solution.* For each fixed realization $(\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_n)$, it is well known that the first-best solution is simply the mean of the ideal points

$$\mathbf{u} = \bar{\mathbf{t}} \equiv \frac{1}{n}\sum_{i=1}^{n}\mathbf{t}_i.$$

Hence, the first-best (per capita) expected utility is the variance (with negative sign)

$$-\frac{1}{n}\sum_{i=1}^{n}\left|\left|\mathbf{t}_i - \bar{\mathbf{t}}\right|\right|^2.$$

In Section 5, we shall extend our analysis to preferences generated by other norms induced by inner products.

2.1 Re-packaging Issues via Rotations

We consider voting by simple majority on each separate dimension. Our focus on simple majority voting stems from its wide applicability and its actual use in practice. We do not a priori restrict the issues on the ballot to be X and Y. Instead, new issues can be created through "re-packaging and bundling" the basic issues X and Y.

We model packaging and bundling of issues through rotations in the plane. Recall that, for fixed Cartesian coordinates, rotating a point $(x, y) \in \mathbb{R}^2$ counter-clockwise by an angle of θ can be represented by the multiplication of the vector (x, y) with a rotation matrix $R(\theta)$. The resulting, rotated vector (z_-, z_+) is given then by

$$\begin{pmatrix} z_- \\ z_+ \end{pmatrix} = \underbrace{\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}}_{R(\theta)} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix}.$$

Equivalently, one can obtain (z_-, z_+) by rotating the original Cartesian coordinates clockwise around the fixed origin by an angle of θ to obtain new orthogonal coordinates, and then projecting (x, y) to the new coordinates.

Let (Z_-, Z_+) denote the new random vector obtained from rotating the random vector (X, Y) by an angle of θ:

$$Z_-(\theta) = X\cos\theta - Y\sin\theta, \quad (1)$$

$$Z_+(\theta) = X\sin\theta + Y\cos\theta. \quad (2)$$

Voters then vote on the new issues Z_- and Z_+, instead of the original issues X and Y.[14] By the simple majority rule, the voting outcome will be $(m_-(\theta, \mathbf{t}_1, ..., \mathbf{t}_n), m_+(\theta, \mathbf{t}_1, ..., \mathbf{t}_n))$ where

$$m_-(\theta, \mathbf{t}_1, ..., \mathbf{t}_n) = \text{median}(x_1\cos\theta - y_1\sin\theta, ..., x_n\cos\theta - y_n\sin\theta), \quad (3)$$

$$m_+(\theta, \mathbf{t}_1, ..., \mathbf{t}_n) = \text{median}(x_1\sin\theta + y_1\cos\theta, ..., x_n\sin\theta + y_n\cos\theta), \quad (4)$$

are the *marginal medians* after the rotation.[15]

It is easy to verify that the mean $\bar{\mathbf{t}}$ of $\mathbf{t}_1, ..., \mathbf{t}_n$ is *rotation equivariant*, i.e. the mean of rotated peaks is simply the rotated mean of the original peaks. In marked contrast, the marginal medians $(m_-(\theta, \mathbf{t}_1, ..., \mathbf{t}_n), m_+(\theta, \mathbf{t}_1, ..., \mathbf{t}_n))$ are **not** rotation equivariant, i.e., rotating and taking medians is not the same as taking medians and rotating. Therefore, rotations are instruments by which the planner may use to influence welfare. To illustrate, consider Figure 1 below with three voters. A, B, and C are voters' ideal points. Original coordinates are drawn in green, rotated coordinates are drawn in red. The green star is the outcome of voting along the original axes (x, y). The red one is the outcome of voting along the rotated axes (x', y'). It is clear that the mean of ideal points is rotation equivariant, the median is not.

[14] We abuse here notation by denoting by the same capital letters both the underlying dimensions (or issues) and the random variables governing the distribution of peaks on those respective dimensions.

[15] Other than marginal median, there are several other multivariate generalizations of univariate median. See Small [1990] for a review of different definitions of multi-dimensional medians and their (lack of) equivariance properties.

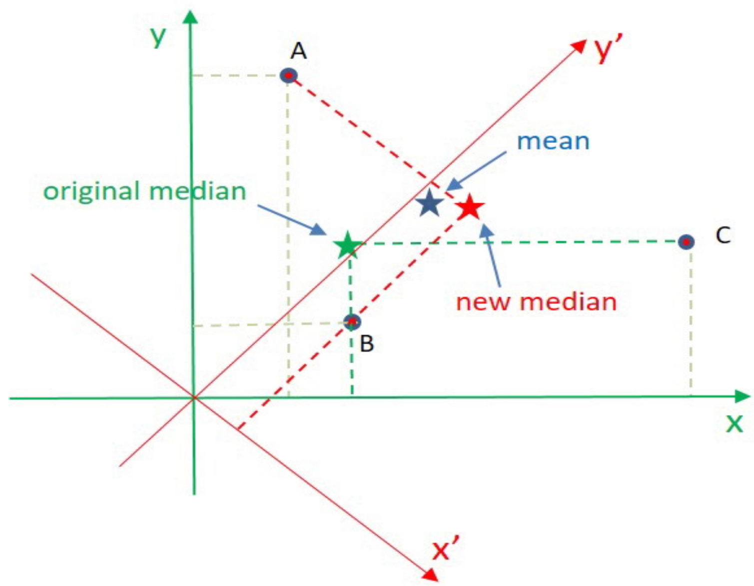

Figure 1: Median is not rotation equivariant.

The reason for this complex behavior is the non-linearity of the median of random variables under convolutions, as illustrated by the following example.

Example 1 *Let (X, Y) denote a random vector on the plane. Suppose that X and Y are I.I.D. exponentially distributed with $f_X(x) = e^{-x}$ for all $x \geq 0$ and $f_Y(y) = e^{-y}$ for all $y \geq 0$. The means are $\mu_X = \mu_Y = 1$ and the medians are $m_X = m_Y = \ln 2$. Rotating the coordinates clockwise by $\frac{\pi}{4}$ and then projecting (X, Y) to the new coordinates, yield a new random vector $(Z_-, Z_+) = (\frac{\sqrt{2}}{2}X - \frac{\sqrt{2}}{2}Y, \frac{\sqrt{2}}{2}X + \frac{\sqrt{2}}{2}Y)$. Z_- is symmetric, so its median and mean are both equal to zero, and the mean of Z_+ equals $\frac{\sqrt{2}}{2}(\mu_X + \mu_Y) = \sqrt{2}$. In contrast, the median of Z_+ is not equal to $\frac{\sqrt{2}}{2}(m_X + m_Y)$, or equivalently, $m_{X+Y} \neq m_X + m_Y$. To see this, note that the density of $X + Y$ is given by*

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(z-t)\, f_Y(t)\, dt = \int_0^z e^{-(z-t)} e^{-t} dt = ze^{-z}, \text{ for all } z \geq 0.$$

Since

$$F_{X+Y}(m_X + m_Y) = \int_0^{2\ln 2} ze^{-z} dz = \frac{3}{4} - \frac{1}{2}\ln 2 \approx 0.4 < F_{X+Y}(m_{X+Y}) = 1/2,$$

it follows that $m_{X+Y} > m_X + m_Y$.

More generally, we could also consider an additional translation of the origin, say by a vector **w,** to obtain new orthogonal coordinates (and thus create new issues).

The joint operation of rotation and translation can also be represented by a linear matrix.[16] But, medians (and means) are translation equivariant, and thus there is no extra welfare advantage from such translations. Therefore, we focus below on the family of rotations - the linear isometries with determinant +1 that fix the origin - described by the angle of rotation θ relative to standard Cartesian coordinates.

2.2 The Set of Voting Mechanisms

For any rotation angle $\theta \in [0, 2\pi]$, we define the direct *marginal median mechanism* φ_θ as

$$\varphi_\theta(\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_n) = (m_-(\theta, \mathbf{t}_1, .., \mathbf{t}_n), m_+(\theta, \mathbf{t}_1, .., \mathbf{t}_n)), \tag{5}$$

where $(m_-(\theta, \mathbf{t}_1, .., \mathbf{t}_n), m_+(\theta, \mathbf{t}_1, .., \mathbf{t}_n))$ is the marginal median with respect to rotation θ and reported peaks $\mathbf{t}_i$ as defined in (3) and (4). Since both rotations and medians are continuous functions, $\varphi_\theta(\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_n)$ is continuous in θ and in all its other arguments.

A direct revelation mechanism $\psi(\mathbf{t}_i, \mathbf{t}_{-i})$ is dominant-strategy incentive compatible (DIC) if, for any voter i, any realizations $\mathbf{t}_i$ and $\mathbf{t}_{-i}$, and any reporting strategy profile $\hat{\mathbf{t}}_{-i}(\mathbf{t}_{-i})$ of other voters, voter i's utility $-\left\|\mathbf{t}_i - \psi\left(\mathbf{t}_i, \hat{\mathbf{t}}_{-i}(\mathbf{t}_{-i})\right)\right\|^2$ is maximized by truthfully revealing his type $\mathbf{t}_i$. It is easily seen that the direct revelation mechanism φ_θ defined in (5) is DIC. Surprisingly, as shown by Kim and Roush [1984] and Peters et al. [1992], the set of marginal median mechanisms (for all possible rotations) coincides with the entire class of anonymous, Pareto efficient and DIC mechanisms.[17] This provides a complementary justification for our focus on simple-majority voting mechanisms.

The mechanism φ_θ can be decentralized (via an indirect mechanism) by first defining the issues (via rotations) and then voting sequentially by simple majority, one issue at a time, using a binary, sequential voting procedure with a convex agenda (such as

[16] This set of general transformation matrices (rotation and translation) is called the *special orthogonal group* for the plane, and is denoted by $SO(2)$. Each matrix in $SO(2)$ is an orthogonal matrix. It is special because the determinant of each matrix is +1, whereas the determinant could be −1 for other orthogonal transformations such as reflections. Rotations form the subgroup that fixes the origin.

[17] A mechanism ψ is *anonymous* if, for any profile of reports $(\mathbf{t}_i, \mathbf{t}_{-i})$, $\psi(\mathbf{t}_1, ..., \mathbf{t}_i, ..., \mathbf{t}_n) = \psi\left(\mathbf{t}_{p(1)}, ..., \mathbf{t}_{p(i)}, ..., \mathbf{t}_{p(n)}\right)$, where p is any permutation of the set $\{1, ..., n\}$. A mechanism ψ is *Pareto efficient* (or Pareto optimal) if, for any profile of reports $(\mathbf{t}_i, \mathbf{t}_{-i})$, there is no alternative $\mathbf{v}$ such that $||\mathbf{t}_i - \mathbf{v}||^2 \leq ||\mathbf{t}_i - \psi(\mathbf{t}_i, \mathbf{t}_{-i})||^2$ for all i, with strict inequality for at least one agent. Note that their characterization fails in higher dimensions because anonymous, Pareto efficient and DIC mechanisms need not exist. Hence, our analysis can be extended to higher dimensional problems, but the solution need not be ex-post Pareto efficient.

those used by all democratic legislatures).[18] The overall outcome does not depend on the order in which the issues are put up to vote, and is the vector of marginal medians $(m_{-}(\theta,\mathbf{t}_1,...,\mathbf{t}_n), m_{+}(\theta,\mathbf{t}_1,...,\mathbf{t}_n))$. This forms an incidence of the structure induced equilibrium à la Shepsle [1979].

Two rotation angles, $\theta = 0$ and $\theta = \pi/4$, are of particular interest and have natural interpretations. When $\theta = 0$, voters are asked to vote on the original issues X and Y. For $\theta = \pi/4$ we have

$$\begin{aligned} m_{-}(\pi/4,\mathbf{t}_1,...,\mathbf{t}_n) &= \frac{\sqrt{2}}{2}\text{median } (x_1 - y_1, ..., x_n - y_n), \\ m_{+}(\pi/4,\mathbf{t}_1,...,\mathbf{t}_n) &= \frac{\sqrt{2}}{2}\text{median } (x_1 + y_1, ..., x_n + y_n). \end{aligned}$$

Therefore, under the $\pi/4$ rotation, the vote is on issues $X + Y$ and $X - Y$, rather than on X and Y. Once voters have decided on $X + Y$ and $X - Y$, the planner can then obviously recover X and Y. The two-step voting procedure associated with the $\pi/4$-rotation resembles the "top-down" budgeting procedure widely used in practice: first a total budget is determined, and then it is allocated among several items. On the other hand, the voting procedure associated with the 0-rotation resembles the "bottom-up" budgeting procedure: agents vote on separate budgets for individual items, and the total budget is gradually obtained as the sum of the individual budgets.

Remark 1 *We focus here on orthogonal coordinates. This is without loss of generality: for any equilibrium outcome obtained by voting along coordinates generated by a non-orthogonal base, there always exists an orthogonal base that yields the same voting outcome. The difference is that under a non-orthogonal base, the order in which the issues are put up to vote does matter. To illustrate, consider the following standard implementation of the* $\pi/4$ *rotation in practice: after the total sum* $(X+Y)$ *was determined, voters are asked to vote on* X *(or on* Y*) rather than on the orthogonal difference* $(X - Y)$*. We show that as long as* $(X + Y)$ *is voted upon first, any issue voted upon at the second stage that is not colinear with* $X + Y$ *will yield the same equilibrium outcome as under voting according to* $(X - Y)$*. To see this, consider the case where voters vote first on* $X + Y$*, and then on* X*, and the second-stage strategy of voter* i *with ideal point* (x_i, y_i)*. The first stage decision imposes then a budget line*

[18] At each stage of convex, sequential procedure on a fixed dimension, a binary decision is collectively taken among two ideologically coherent sets of alternatives that create a clear left-right divide. For details see Gershkov, Moldovanu and Shi [2017] and Kleiner and Moldovanu [2017].

(the purple dash line in Figure 2) on which the final voting outcome must lie.

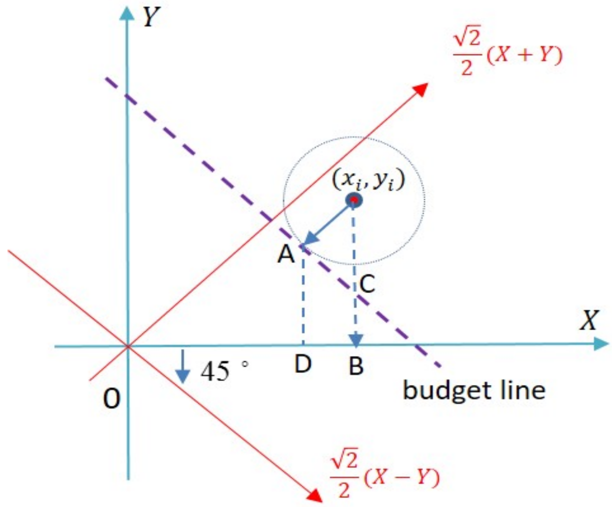

Figure 2. Alternative implementation of the top-down procedure.

Let A and B denote the points obtained by projecting (x_i, y_i) on the budget line, and on the X axis, respectively, and let D denote the projection A to the X axis. Then, at the second stage voting on X, voter i's dominant strategy is to vote for point D rather than point B : whenever i is pivotal, voting D yields point A on the budget line, which is closest to his ideal point. On the other hand, A is exactly the point that i would have voted for if the second stage vote were on the difference $X - Y$. Note that the above argument is independent of the number of voters and can be easily generalized to other non-orthogonal bases.

3 The Limit Case when the Number of Agents Is Large

The full probabilistic optimization problem can be rewritten as

$$(\mathcal{P}_0) \qquad \min_{\theta\in[0,2\pi]} \int_D ... \int_D \left(\frac{1}{n} \sum_{i=1}^{n} ||R(\theta)\,\mathbf{t}_i - \varphi_\theta(\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_n)||^2 \right) f(\mathbf{t}_1)...f(\mathbf{t}_n) d\mathbf{t}_1...d\mathbf{t}_n.$$

We focus here on the solution to problem $(\mathcal{P}_0)$ when the number of agents is large. The resulting optimal mechanism will be incentive compatible, Pareto efficient and anonymous for any (odd) number of voters. For I.I.D. random variables $\{X_i\}_{i=1}^{\infty}$ with finite mean μ_X and variance σ_X^2, we know from the central limit theorem that

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu_X\right) \to N(0, \sigma_X^2).$$

Bahadur (1966) shows that the quantiles of large samples display a similar behavior. In particular,

$$\sqrt{n}(X_{(n+1)/2:n} - m_X) \rightarrow N\left(0, \frac{1}{4f^2(m_X)}\right),$$

where $X_{(n+1)/2:n}$ is the median order statistic, and where m_X is the median of the distribution. Thus, as n goes to infinity, the sample median converges to the median of the underlying distribution and, of course, the sample mean converges to the mean.

By applying the above limit results to our setting, we obtain that, as $n \rightarrow \infty$,

$$\begin{pmatrix} m_-(\theta, \mathbf{t}_1, .., \mathbf{t}_n) \\ m_+(\theta, \mathbf{t}_1, .., \mathbf{t}_n) \end{pmatrix} \longrightarrow \begin{pmatrix} m_-(\theta) \\ m_+(\theta) \end{pmatrix} \equiv \begin{pmatrix} \text{median } (X\cos\theta - Y\sin\theta) \\ \text{median } (X\sin\theta + Y\cos\theta) \end{pmatrix}.$$

Furthermore, since the norm $||\cdot||$ is continuous, we obtain that, as $n \rightarrow \infty$,

$$\begin{aligned} & \frac{1}{n}\sum_{i=1}^{n} ||R(\theta)\mathbf{t}_i - \varphi_\theta(\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_n)||^2 \\ = \ & \frac{1}{n}\sum_{i=1}^{n} \left[(x_i\cos\theta - y_i\sin\theta - m_-(\theta, \mathbf{t}_1, .., \mathbf{t}_n))^2 + (x_i\sin\theta + y_i\cos\theta - m_+(\theta, \mathbf{t}_1, .., \mathbf{t}_n))^2\right] \\ \rightarrow \ & \mathbb{E}\,||X\cos\theta - Y\sin\theta - m_-(\theta), X\sin\theta + Y\cos\theta - m_+(\theta)||^2 \\ = \ & \sigma_X^2 + \sigma_Y^2 + \left(\mu_-(\theta) - m_-(\theta)\right)^2 + \left(\mu_+(\theta) - m_+(\theta)\right)^2, \end{aligned}$$

where the two coordinates of the rotated mean are

$$\mu_-(\theta) = \mu_X\cos\theta - \mu_Y\sin\theta, \text{ and } \mu_+(\theta) = \mu_X\sin\theta + \mu_Y\cos\theta.$$

Therefore, in the limit where n is very large, the problem becomes

$$(\mathcal{P}_1) \quad \min_{\theta\in[0,2\pi]} \left(\mu_-(\theta) - m_-(\theta)\right)^2 + \left(\mu_+(\theta) - m_+(\theta)\right)^2 + \sigma_X^2 + \sigma_Y^2.$$

In other words, we look for the rotation that creates the marginal median vector with the minimum distance from the mean.

For most parts of the analysis below, it will be convenient to normalize the means of X and Y to be zero – such a normalization is without loss of generality because of the translational equivariance of both mean and median. Let us define the normalized random variables $\tilde{X}$ and $\tilde{Y}$ as

$$\tilde{X} = X - \mu_X \text{ and } \tilde{Y} = Y - \mu_Y.$$

The corresponding normalized marginal medians $(\tilde{m}_-(\theta), \tilde{m}_+(\theta))$ are

$$\tilde{m}_-(\theta) = m_-(\theta) - \mu_-(\theta) \quad \text{and} \quad \tilde{m}_+(\theta) = m_+(\theta) - \mu_+(\theta).$$

We further note that it is without loss of generality to restrict attention to rotations in the interval $[0, \pi/2]$. That is because, for any $\theta \in [\pi/2, 2\pi]$ that minimizes the planner's objective, there exists $\theta' \in [0, \pi/2]$ that attains the same minimum.[19] Hence, the planner's problem can be rewritten as

$$(\mathcal{P}_2) \qquad \min_{\theta \in [0,\pi/2]} \tilde{m}_-^2(\theta) + \tilde{m}_+^2(\theta) + \sigma_X^2 + \sigma_Y^2.$$

Since variances are fixed, the planner's goal under this normalization is simply to find the rotation resulting in a marginal median vector with minimum norm. To simplify notation, we shall drop the tilde symbol for normalized random variables where no confusion can arise.

Remark 2 *We would like to comment here on the feasibility of the first-best solution.*

1. *With a continuum of voters, the planner could, in principle, dictate the mean as the collective choice without seeking any input from the voters. But, this would require detailed knowledge about the joint distribution of individuals' preferences. In contrast, voting by simple majority in each dimension is practical and indeed often observed in reality because it is always incentive compatible, and because its execution does not require any prior knowledge about the distribution. None of our theorems or propositions (e.g., Theorems 1-3, Propositions 1-3) requires the planner to know the exact distribution: it is sufficient to know that the joint distribution belongs to a broad class.*

2. *If the number of voters is finite, the first-best solution, defined as the sample mean of the voters' ideal points, is not implementable because each agent can advantageously move the mean towards her ideal point by reporting a false peak. The individual influence on the mean is unbounded (unless the distribution of peaks is on a bounded interval). Thus, even if the number of voters is large, the possibility to tilt the mean in one's favor may still be substantial.*

3.1 Sub-Optimality of Voting on Independent Issues

In this subsection, we assume that the unrotated marginals X and Y are independent. We work on the normalized version of the planner's problem $(\mathcal{P}_2)$ and show that the zero-rotation yields a local maximum of the norm of the normalized marginal median, i.e., it leads to a local utility minimum.

[19] This claim is a direct consequence of simple trigonometric identities, and we omit the proof.

Theorem 1 *Assume that X and Y are independent. The rotation with angle $\theta = 0$ is a local utility minimum if*

$$m_X f'_X(m_X) \geq 0, m_Y f'_Y(m_Y) \geq 0, m_X^2 + m_Y^2 \neq 0. \tag{6}$$

Proof. See Appendix A. ■

If random variables X and Y are unimodal, then the rotation of $\theta = 0$ is a local utility minimum if the median lies between the mode and the mean.[20] This alternative sufficient condition is simple and intuitive: there are elegant, general characterizations of distributions where such orders of the mode, median and mean hold (see for example, Dharmadhikari and Joag-Dev [1988], Basu and DasGupta [1997]).

Corollary 1 *Assume that X and Y are independent and $m_X^2 + m_Y^2 \neq 0$. Suppose that X and Y are unimodal and satisfy*

$$\begin{aligned} M_X &\leq m_X \leq \mu_X \quad \text{or} \quad \mu_X \leq m_X \leq M_X \\ M_Y &\leq m_Y \leq \mu_Y \quad \text{or} \quad \mu_Y \leq m_Y \leq M_Y \end{aligned}$$

where M, m, μ are mode, median and mean, respectively. Then the rotation with angle $\theta = 0$ is a local utility minimum.

Proof. If $M_X \leq m_X \leq \mu_X = 0$ (where the last equality holds by normalization), then $m_X \leq 0$ and $f'(m_X) \leq 0$ because m_X is to the right of the mode. Hence $m_X f'_X(m_X) \geq 0$. If $0 = \mu_X \leq m_X \leq M_X$, then $m_X \geq 0$ and $f'(m_X) \geq 0$ because m_X is to the left of the mode. Hence $m_X f'_X(m_X) \geq 0$, and analogously for Y. ■

The proof of Theorem 1 proceeds as follows: the rotation $\theta = 0$ yields a local maximum of the norm of the normalized marginal median if it is a critical point

$$m_-(0)m'_-(0) + m_+(0)m'_+(0) = 0, \tag{7}$$

and if it satisfies the following local second-order condition

$$m''_-(0)m_-(0) + (m'_-(0))^2 + m''_+(0)m_+(0) + (m'_+(0))^2 < 0. \tag{8}$$

The proof verifies that $m'_-(0) = m'_+(0) = 0$ (so condition (7) is trivially satisfied), and that condition (6) in Theorem 1 implies condition (8).

[20] A random variable Z is *unimodal* if its density $f(z)$ has a single mode (or peak).

The geometric intuition of the sub-optimality of voting on independent issues is illustrated in Figure 3 below:

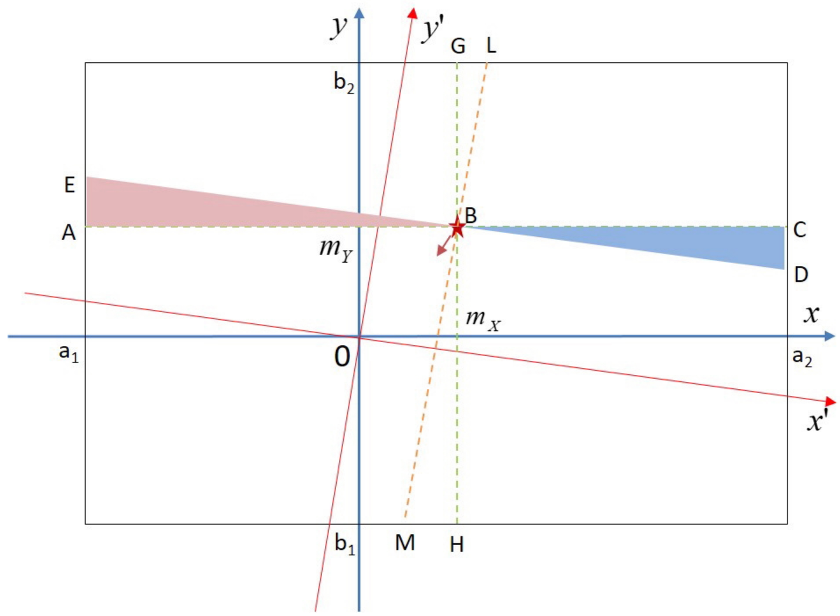

Figure 3. Small rotation improves welfare.

Assume that $0 = \mu_X \leq m_X$ and $0 = \mu_Y \leq m_Y$. We want to show that a small rotation improves welfare if $f'_X(m_X) \geq 0$ and $f'_Y(m_Y) \geq 0$. Assume that the unrotated median is B. Therefore, by independence, there is a mass of 50% above the AC line and a mass of 50% to the right of GH line. Consider a small rotation with angle $\theta > 0$, so that new axes are x' and y'. We want to show that this shifts the new median towards the mean $(0, 0)$,i.e., that the median moves towards the south-west. Consider the projection of B onto the new, rotated axes: the result obtains if the mass above DE and the mass to the right of LM are both below 50%. If the area of ABE is larger than the one of BCD, we obtain that the mass above ED is indeed smaller than 0.5 (the comparison for the other dimension is analogous).

For illustration purpose, let us assume that X and Y distribute on bounded intervals $[a_1, a_2]$ and $[b_1, b_2]$, respectively. The line DE passing through point B is given by $y = m_Y + (m_X - x)\tan\theta$. Therefore, the difference between the areas ABE and BCD is

$$ABE - BCD = \int_{a_1}^{a_2} [F_Y(m_Y + (m_X - x)\tan\theta) - F_Y(m_Y)] f_X(x)\,dx.$$

Since $f'_Y(m_Y) \geq 0$, F_Y is locally convex at m_Y. Therefore, for sufficiently small θ, the curve $F_Y(m_Y + (m_X - x)\tan\theta)$ with $x \in [a_1, a_2]$ lies above the tangent line $F_Y(m_Y) + f_Y(m_Y)(m_X - x)\tan\theta$. That is,

$$F_Y(m_Y + (m_X - x)\tan\theta) \geq F_Y(m_Y) + f_Y(m_Y)(m_X - x)\tan\theta.$$

As a result, for sufficiently small θ, we have

$$ABE - BCD \geq \int_{a_1}^{a_2} f_Y(m_Y)(m_X - x)\tan\theta f_X(x)\,dx = f_Y(m_Y)\, m_X \tan\theta > 0,$$

as desired. The argument for the other dimension is analogous.

Intuitively, area ABE represents voters who have their preferred y coordinate marginally above m_Y and who, after rotation, would switch their support from alternatives above the line $y = m_Y$ to alternatives below the line $y' = m_Y$. In contrast, area BCD represents voters who have their y coordinate marginally below m_Y and who would switch their support from alternatives below the line $y = m_Y$ to alternatives above the line $y' = m_Y$. Since $F(y)$ is locally convex at $y = m_Y$, there are more voters in area ABE than in BCD, and thus more than half of them will vote for alternatives below the line $y' = m_Y$. That is, the median after the rotation will be closer to the origin (the first best).

3.2 The $\pi/4$-Rotation

In this subsection, we assume that X and Y are identically (but not necessarily independently) distributed. By symmetry,

$$m_-(\pi/4) = \text{median}(\frac{\sqrt{2}}{2}(X - Y)) = 0 = \mu_-(\pi/4),$$

and

$$m_+(\pi/4) = \text{median}\,(\frac{\sqrt{2}}{2}(X + Y)) = \frac{\sqrt{2}}{2}\text{median}\,(X + Y).$$

Hence, the $\pi/4$-rotation is a natural candidate for improving welfare. It completely eliminates the conflict arising between efficiency and incentive compatibility along one dimension – all remaining conflict is concentrated in the other dimension, as illustrated in the following figure (assuming $m_X > \mu_X = 0$) where (m_X, m_Y) is the unrotated median and the red star is the $\pi/4$-rotated median $m_+(\pi/4)$:

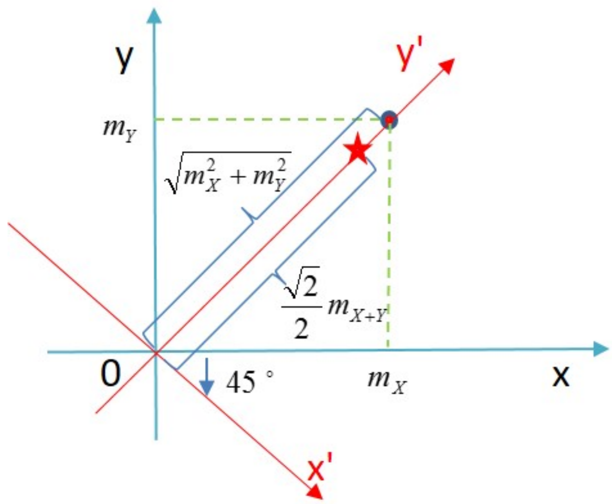

Figure 4. The $\pi/4$-rotation with symmetric marginals.

Proposition 1 *Suppose that X and Y are I.I.D., and the density f_X satisfies the following regularity condition:*

$$\lim_{x\to\infty} f_X\left(\sqrt{2}m_+(\pi/4) - x\right) f_X(x)\left(2x - \sqrt{2}m_+(\pi/4)\right)^2 = 0,$$
$$\lim_{x\to-\infty} f_X\left(\sqrt{2}m_+(\pi/4) - x\right) f_X(x)\left(2x - \sqrt{2}m_+(\pi/4)\right)^2 = 0.$$

Then $\theta = \pi/4$ is a critical point, i.e., it satisfies the first order condition.

Proof. See Appendix A. ■

The above regularity condition is satisfied if the distribution has a bounded support or a thin tail. If we could verify second-order conditions either locally or globally, then Proposition 1 could tell us whether $\theta = \pi/4$ is a local or global utility maximum. Unfortunately, the second order conditions, evaluated at $\theta = \pi/4$, turn out to be very elusive.

Our next result offers sufficient conditions for the optimality of the $\pi/4$-rotation. It requires the following definition.

Definition 1 *A vector (a, b) is said to majorize (a', b'), written as $(a, b) \succ (a', b')$, if $a + b = a' + b'$ and if $\max(a, b) \geq \max\{a', b'\}$. A function $h(a, b)$ is said to be Schur-convex (concave) in (a, b) if $h(a'', b'') \geq (\leq) h(a', b')$ whenever $(a'', b'') \succ (a', b')$.*

Proposition 2 *Suppose that X and Y are identically distributed and $m_X \neq \mu_X$. The $\pi/4$-rotation attains the welfare maximum if either*

1. $m_+(\theta) < \mu_+(\theta)$ *for all* $\theta \in \left[0, \frac{\pi}{4}\right]$*, and the function*

$$\Pr\left(X \sin\theta + Y \cos\theta \le z\right)$$

is Schur-concave in $(\sin^2\theta, \cos^2\theta)$ *for all* $\theta \in \left[0, \frac{\pi}{4}\right]$ *and all* $z \in [m_X, 0]$*;*

or

2. $m_+(\theta) > \mu_+(\theta)$ *for all* $\theta \in \left[0, \frac{\pi}{4}\right]$*, and the function*

$$\Pr\left(X \sin\theta + Y \cos\theta \le z\right)$$

is Schur-convex in $(\sin^2\theta, \cos^2\theta)$ *for all* $\theta \in \left[0, \frac{\pi}{4}\right]$ *and* $z \in [0, m_X]$*.*

Proof. See Appendix A. ■

If $\Pr\left(X \sin\theta + Y \cos\theta \le z\right)$ is Schur-concave for all $\theta \in \left[0, \frac{\pi}{4}\right]$, and if the rotated median is always below the mean, it must hold that

$$m_X \le m_{X \sin\theta + Y\cos\theta} \le m_{\frac{\sqrt{2}}{2}X + \frac{\sqrt{2}}{2}Y} \le \mu_X.$$

Hence, the distance between the mean and the rotated median $m_{X \sin\theta + Y\cos\theta}$ is smallest when $\theta = \pi/4$. The sufficient conditions in Proposition 2 only involve the model's primitives (i.e., the distributions of types) and can be, in principle, checked for any distribution.[21]

For example, we verified that $\Pr\left(X \sin\theta + Y \cos\theta \le z\right)$ is Schur-concave if X and Y are I.I.D. exponential and thus the $\pi/4$-rotation is globally optimal in that case.[22] For other standard distributions such as gamma, Pareto and Rayleigh, we used Mathematica to plot the aggregate expected welfare as a function of the rotation angle $\theta \in [0, \pi/2]$. Our simulations suggest that the $\pi/4$-rotation is optimal for these distributions, but we were unable to analytically prove it. In general, the $\pi/4$-rotation may not be optimal, as illustrated by Example 2 below. Therefore, some restrictions on the symmetric marginals are indeed necessary for the optimality of the $\pi/4$-rotation.

Example 2 *Let* (X, Y) *denote a random vector on the plane. Suppose that* X *and* Y *are I.I.D. according to the following discrete distribution:*

values of X	0	0.45	1
probability	0.4	0.3	0.3

[21] Similar Schur-concavity/convexity conditions appear in the literature: For example, if X, Y are **non-negative** I.I.D. random variables with a log-concave density then $\Pr(aX + bY \le z)$ is known to be Schur-concave function of (a^2, b^2) for all z (see Karlin and Rinott [1983]). We cannot directly use this result because of the non-negativity restriction.

[22] The verification details for the exponential distribution are available upon request.

so that $\mu_X = \mu_Y = 0.435$ *and* $m_X = m_Y = 0.45$*. The distribution of* $X+Y$ *is given by*

values of $X+Y$	0	0.45	0.90	1	1.45	2
probability	0.16	0.24	0.09	0.24	0.18	0.09

so that $\mu_{X+Y} = 0.87$ *and* $m_{X+Y} = 1$*. The expected utility from the* 0*-rotation is*

$$-2\left(\mu_X - m_X\right)^2 = -2\left(0.435 - 0.45\right)^2 = -0.000\,45$$

and the expected utility from the $\pi/4$*-rotation is*

$$-\left(\frac{\sqrt{2}}{2}\left(\mu_{X+Y} - m_{X+Y}\right)\right)^2 = -\left(\frac{\sqrt{2}}{2}\left(0.87-1\right)\right)^2 = -0.008\,45$$

Therefore, the $\pi/4$*-rotation is strictly dominated by the* 0*-rotation. Since the welfare dominance is strict, we can approximate the discrete distribution by a continuous distribution and maintain it.*

3.3 When does "Top-Down" Dominate "Bottom-Up"?

We now compare the expected utility under the $\pi/4$-rotation with that under the 0-rotation. As is apparent from Figure 4, this amounts to check whether the original coordinate-wise median vector (m_X, m_Y) is closer to the origin than the new coordinate-wise median vector $(m_{X+Y}/2, m_{X+Y}/2)$. Therefore, if $m_X < \mu_X$ and $m_X + m_Y < m_{X+Y}$, or if $m_X > \mu_X$ and $m_X + m_Y > m_{X+Y}$, then the $\pi/4$-rotation dominates the zero-rotation.

Assuming that X and Y are I.I.D., we present below a simple sufficient condition that simultaneously guarantees $m_X < (>)\, \mu_X$ and $m_X + m_Y < (>)\, m_{X+Y}$.[23] The need to control for sub/super-additivity of medians parallel the conditions on second-highest order statistics for bundling in auctions (see Palfrey [1983]).

Proposition 3 *Suppose that* X *and* Y *are I.I.D. and that* $m_X \neq \mu_X$*. The expected utility at* $\theta = \frac{\pi}{4}$ *exceeds the expected utility at* $\theta = 0$ *if either*

$$F_X\left(m_X + \varepsilon\right) + F_X\left(m_X - \varepsilon\right) \leq 1 \text{ for all } \varepsilon > 0, \quad (9)$$

[23] As is illustrated in Example 1, both condition $m_X < \mu_X$ and the super-additivity condition $m_X + m_Y < m_{X+Y}$ hold for the exponential distribution which is strictly concave. We show in Section 7.5 of Appendix A that the super-additivity condition is satisfied for the gamma distribution (a generalization of the exponential) and the Rayleigh distribution, where the sufficient condition (9) may not be easily checked, or does not hold. There we also construct, by using a copula, an example where independence is not necessary for the $\pi/4$-rotation to dominate the 0-rotation.

or

$$F_X(m_X + \varepsilon) + F_X(m_X - \varepsilon) \geq 1 \text{ for all } \varepsilon > 0. \tag{10}$$

In particular, condition (9) implies $m_X < \mu_X$ and $m_X + m_Y < m_{X+Y}$, and is satisfied if F_X is strictly concave. Condition (10) implies $m_X > \mu_X$ and $m_X + m_Y > m_{X+Y}$, and is satisfied if F_X is strictly convex.

Proof. See Appendix A. ■

It is worth noting that van Zwet [1979] shows that condition (9) implies $\mu_X > m_X > M_X$ and (10) implies $\mu_X < m_X < M_X$. It follows from Corollary 1 that each of the two conditions is also sufficient for the zero-rotation to be sub-optimal.

Remark 3 *Whenever the median function is super (sub)-additive, the top-down procedure where a total budget is determined first leads to a higher (lower) overall budget than the bottom-up procedure where votes are item-by-item and where the total budget is gradually determined.*[24]

4 Bounds on Relative Efficiency

In this section we provide several lower bounds on the (relative) efficiency loss of the marginal median mechanisms augmented by rotations. We keep the assumption that the number of agents is large. The various bounds are obtained under different distributional assumptions governing the distribution of voter's ideal points, and the proofs use several classical statistical inequalities, and some more recent concentration inequalities. In particular, for the logconcave case studied by Caplin and Nalebuff ([1988], [1991]), the lower bound is 88% of the first-best utility.

Note that each assertion in the following Theorem holds for a large class of distributions, and therefore that the results do not require exact knowledge of the particular distribution (as long as it is known that it belongs to the respective class). In particular, the optimal rotation achieves, in each case, a possibly higher relative efficiency.

Assume that ideal points are distributed such that the marginals are given by random variables (X, Y) where X and Y are not necessarily identical, and are potentially correlated. Since the results heavily use statistical results that establish relations between the mean, median and variance, we work here with the **non-normalized** variables (so that the role of the mean and its relations to the other statistics does not get obscured by the normalization we used above). The first-best expected utility,

[24] Note that this question is not identical to the question of utility comparisons.

attained by choosing the mean in each coordinate, decreases as variances increase and is given by

$$-\mathbb{E}(X-\mu_X)^2 - \mathbb{E}(Y-\mu_Y)^2 = -\sigma_X^2 - \sigma_Y^2.$$

The expected utility of rotated medians with angle θ is given by

$$U(\theta) = -\sigma_X^2 - \sigma_Y^2 - \left(\mu_-(\theta) - m_-(\theta)\right)^2 - \left(\mu_+(\theta) - m_+(\theta)\right)^2.$$

Thus, the relative efficiency of the rotation with angle θ is given by:

$$EF(\theta) = \frac{\sigma_X^2 + \sigma_Y^2}{\sigma_X^2 + \sigma_Y^2 + \left(\mu_-(\theta) - m_-(\theta)\right)^2 + \left(\mu_+(\theta) - m_+(\theta)\right)^2} \leq 1.$$

Two forces play here a role: on the one hand, a distribution that is concentrated around a central location (such as the mean or the median) will have a small difference between mean and median, which tends to increase the relative efficiency. On the other hand, such a distribution also has a low variance so that the difference between mean and median plays a bigger overall role.[25] The first-best outcome can be attained by majority voting (in the limit with a large number of agents) if the distributions of both X and Y are symmetric around their respective means (e.g., both are normally distributed). In this case we have $\mu_-(\theta) = m_-(\theta)$ and $\mu_+(\theta) = m_+(\theta)$.

A random variable X has *increasing failure rate* (IFR) if its hazard rate $f(x)/(1-F(x))$ is increasing in x.

Theorem 2 *The following relative efficiency bounds hold:*

1. *For any random variables X and Y and for any angle θ, $EF(\theta) \geq \frac{1}{2}$.*

2. *For any unimodal random variables X and Y and for any angle θ, $EF(\theta) > \frac{5}{8}$.*

3. *For any random variables X and Y that have an increasing failure rate (IFR) and that satisfy $\mu_X \leq m_X$ and $\mu_Y \leq m_Y$, and for any angle θ, $EF(\theta) > 0.603$. In addition, if X and Y are I.I.D., then $EF(\frac{\pi}{4}) \geq 0.753$.*

4. *For any X and Y that are identically distributed and for any angle θ, $EF(\frac{\pi}{4}) \geq \frac{2\sigma_X^2}{3\sigma_X^2 + Cov(X,Y)}$. Thus, when X and Y are independent, $EF(\frac{\pi}{4}) \geq \frac{2}{3}$. In the polar, co-monotonic scenario, $EF(\frac{\pi}{4}) = EF(0) \geq \frac{1}{2}$ and welfare cannot be improved by rotation.*[26]

[25] It is interesting to note that the covariance of X and Y does not play a direct role in the efficiency calculations: it only enters in the way medians of convolutions are calculated.

[26] A random vector is *co-monotonic* if and only if it agrees in distribution with a random vector where all components are non-decreasing functions (or all are non-increasing functions) of the same random variable.

5. *If X and Y are I.I.D. and if each has a log-concave density, then $EF(\frac{\pi}{4}) \geq 0.876$.*

Proof. 1. A classical inequality due to Hotelling and Solomons [1932] says that the squared distance between the mean and median of **any** random variable is always less than the variance:

$$(\mu - m)^2 \leq \sigma^2.$$

Therefore,

$$\begin{aligned}\left(\mu_-(\theta) - m_-(\theta)\right)^2 &\leq \sigma_-^2(\theta) = \sigma_X^2 \cos^2\theta + \sigma_Y^2 \sin^2\theta - 2\sin\theta\cos\theta Cov(X,Y), \\ \left(\mu_+(\theta) - m_+(\theta)\right)^2 &\leq \sigma_+^2(\theta) = \sigma_X^2 \sin^2\theta + \sigma_Y^2 \cos^2\theta + 2\sin\theta\cos\theta Cov(X,Y).\end{aligned}$$

We obtain the universal bound:

$$EF(\theta) \geq \frac{\sigma_X^2 + \sigma_Y^2}{2\sigma_X^2 + 2\sigma_Y^2} = \frac{1}{2}.$$

2. For the class of unimodal distributions the squared distance between mean and median is at most $\frac{3}{5}$ variance (see Basu and DasGupta [1997]). Thus, for such distributions we get:

$$EF(\theta) \geq \frac{\sigma_X^2 + \sigma_Y^2}{(\sigma_X^2 + \sigma_Y^2) + \frac{3}{5}(\sigma_X^2 + \sigma_Y^2)} = \frac{5}{8}.$$

3. For the class of distributions with an increasing failure rate (IFR), if $\mu_X \leq m_X$, then we obtain from Rychlik [2000] that

$$\frac{(\mu_X - m_X)^2}{\sigma^2} \leq \frac{(-\log(\frac{1}{2}) - \frac{1}{2})^2}{\frac{3}{4} + \log(\frac{1}{2})} = 0.656,$$

and hence an efficiency rate of

$$EF(\theta) \geq \frac{\sigma_X^2 + \sigma_Y^2}{(\sigma_X^2 + \sigma_Y^2) + 0.656(\sigma_X^2 + \sigma_Y^2)} = \frac{1}{1 + 0.656} = 0.603.$$

If in addition, X and Y are I.I.D., then the convolution of two such variables is again IFR (see Barlow and Proschan [1965]) and we obtain

$$EF(\frac{\pi}{4}) \geq \frac{2\sigma_X^2}{2\sigma_X^2 + 0.656\sigma_X^2} = 0.753.$$

4. If X distributes as Y (not necessarily independent), we know that $X - Y$ is symmetric and hence that $m_-\left(\frac{\pi}{4}\right) = \mu_-\left(\frac{\pi}{4}\right) = 0$. This yields:

$$EF(\frac{\pi}{4}) = \frac{2\sigma_X^2}{2\sigma_X^2 + \left(\mu_+\left(\frac{\pi}{4}\right) - m_+\left(\frac{\pi}{4}\right)\right)^2} \geq \frac{2\sigma_X^2}{3\sigma_X^2 + Cov(X,Y)}$$

Assume that (X_1, Y_1) and (X_2, Y_2) belong to the same Frechet class $M(F_1, F_2)$ of bivariate distributions with fixed marginals F_1 and F_2. Moreover, assume that (X_1, Y_1) $\leq_{PQD}$ (X_2, Y_2) where PQD stands for the *positive quadrant order* (see Lehmann [1966]). This stochastic order measures the amount of positive dependence of the underlying random vectors.[27] We obtain that all one-dimensional variances are identical, but that $Cov(X_1, Y_1) \leq Cov(X_2, Y_2)$. Thus, the worst case efficiency bound is higher when the variates are less positive dependent. In particular, for given marginals, the highest worst-case efficiency of the $\frac{\pi}{4}$ rotation is achieved for the I.I.D. case where $Cov(X, Y) = 0$, and where:

$$EF(\frac{\pi}{4}) \geq \frac{2\sigma_X^2}{3\sigma_X^2} = \frac{2}{3}.$$

The polar case to independence is the case where X and Y are *co-monotonic.* Then their covariance is maximized for given marginals, and their convolution is quantile-additive (see Kaas et al. [2002]). In other words, quantiles and thus medians (i.e., the 50% quantile) are linear functions. Hence we obtain for the median that $m_+(\frac{\pi}{4}) = \sqrt{2}m_X$. Hence,

$$\left(\mu_+\left(\frac{\pi}{4}\right) - m_+\left(\frac{\pi}{4}\right)\right)^2 = (\mu_X - m_X)^2 \leq 2\sigma_X^2$$

and we obtain

$$EF(\frac{\pi}{4}) = EF(0) \geq \frac{1}{2}.$$

This holds analogously for any other rotation.

5. Consider now the I.I.D. case with log-concave densities.[28] Then X and Y are unimodal. Their convolution is log-concave (Prekopa [1973]), and hence also unimodal.[29] Let $f_X = f_Y$ denote the respective log-concave densities. Bobkov and Ledoux [2014] prove that:[30]

$$\frac{1}{12\sigma_X^2} \leq f_X^2(m_X) \leq \frac{1}{2\sigma_X^2}.$$

On the other hand, Ball and Böröczky [2010] prove that:

$$f_X(m_X) \cdot |m_X - \mu_X| \leq \ln\left(\sqrt{\frac{e}{2}}\right).$$

[27]It is implied, for example, by the supermodular order.

[28]Note that any log-concave distribution on the plane yields log-concave marginals (Prekopa [1973]).

[29]The convolution of unimodal densities need not be unimodal. But the convolution of X and Y is unimodal for any Y if and only if X is log-concave (see Ibragimov [1956]).

[30]Interestingly enough, the left hand side of the inequality applies to **any** probabiliy density on the real line.

Combining the two inequalities above yields

$$(m_X - \mu_X)^2 \leq \frac{1}{f_X^2(m_X)} \ln^2\left(\sqrt{\frac{e}{2}}\right) \leq 12\sigma_X^2 \ln^2\left(\sqrt{\frac{e}{2}}\right).$$

The efficiency bound in the log-concave case becomes then

$$EF(\frac{\pi}{4}) \geq \frac{2\sigma_X^2}{2\sigma_X^2 + 12\sigma_X^2 \ln^2\left(\sqrt{\frac{e}{2}}\right)} = \frac{1}{1 + 6\ln^2\left(\sqrt{\frac{e}{2}}\right)} = 0.876.$$

■

The above calculations also show that the improvement obtained by rotation may be significant. Just to give one example, consider the original (e.g., unrotated) distributions for which the Hotelling-Solomons bound is achieved with equality.[31] Then, the welfare in the I.I.D. case without rotation is exactly half of the first-best welfare, while the welfare following the 45 degree rotation is at least two-thirds of the original first best, yielding an improvement of at least 30%.

In Appendix B, we show how the above bounds can be obtained for the case of more dimensions. For example, in the I.I.D, case, the relative efficiency tends to 1 when the number of dimensions becomes infinite.

5 Extension to Other Utility Functions

In this section we briefly illustrate how our method can be applied to a more general class of utility functions that are based on a distance generated by an inner product. Thus, we assume that the utility of agent i with peak $\mathbf{t}_i$ from decision $\mathbf{v} \in D \subseteq \mathbb{R}^2$ is given by

$$-\Delta\left(||\mathbf{t}_i - \mathbf{v}||_I\right),$$

where $||\cdot||_I$ is some inner-product norm, and where Δ is a strictly monotonically increasing function.

Since inner-product norms are strictly convex, choosing a marginal median with respect to any orthogonal coordinates yields a DIC mechanism (see Peters et al. [1993]).[32] Recall that two vectors are orthogonal if their inner-product (that induces the distance function) is zero.

[31] This is a discrete distribution concentrated on two points. But, it can be easily approximated by continuous distribution that satisfy the bound with almost equality, for any needed degree of precision.

[32] These authors also show that, as in the case of the Euclidean norm in the plane, the class of marginal medians coincides with the class of DIC, anonymous and Pareto efficient mechanims.

For the Euclidean norm, every rotation is an isometry that fixes the origin and preserves orthogonality and orientation: it transforms a basis of orthogonal vectors into another such basis, and each oriented orthogonal basis is obtained (modulo translation) from another via a suitable rotation.

In order to proceed in an analogous fashion, we need to first identify the set of isometries: for any inner product norm $||\cdot||_I$ this is always an infinite multiplicative group (see Garcia-Roig [1997]). Because medians and welfare measures that are based on distances are translation equivariant, it is enough, as above, to characterize the sub-group of isometries that fix the origin and that preserve orientation (i.e., their corresponding matrices have determinant +1). We start with the simplest case.

5.1 Weighted Euclidean Norm

An agent with ideal point $\mathbf{t}_i = (x_i, y_i)$ has a weighted Euclidean preference over points $\mathbf{v} = (x, y)$ given by

$$-\beta^2 (x - x_i)^2 - (y - y_i)^2 ,$$

with $\beta > 0$. Note that, without loss of generality, we can always normalize one of the weights to be +1 without changing the underlying (ordinal) preferences. Let

$$M = \begin{pmatrix} \beta^2 & 0 \\ 0 & 1 \end{pmatrix},$$

and define an inner-product and its associated norm by:

$$\begin{aligned} \langle (x_1, y_1), (x_2, y_2) \rangle &\equiv (x_1, y_1) M (x_2, y_2)^T, \\ ||(x, y)|| &\equiv \sqrt{(x, y) M (x, y)^T} = \sqrt{\beta^2 x^2 + y^2}. \end{aligned}$$

The "unit circle" is here an ellipse

$$\beta^2 x^2 + y^2 = 1,$$

with axes parallel to the standard Cartesian coordinate axes. Isometries that fix the origin leave this ellipse invariant (i.e., a point on the ellipse is translated to another point on the ellipse) and can be represented by generalized "rotation" matrices of the form

$$R_\beta(\theta) = \begin{pmatrix} \cos\theta & -\frac{1}{\beta}\sin\theta \\ \beta\sin\theta & \cos\theta \end{pmatrix}.$$

While the mean in each coordinate is still the first-best, the welfare measure changes to incorporate the weight β. By normalizing the mean to zero, the welfare maximization problem becomes:

$$\min_\theta [\beta^2 m_{\beta-}^2 (\theta) + m_{\beta+}^2 (\theta) + \beta^2 \sigma_X^2 + \sigma_Y^2] \Leftrightarrow \min_\theta [\beta^2 m_{\beta-}^2 (\theta) + m_{\beta+}^2 (\theta)],$$

where

$$m_{\beta-}(\theta) = \text{median } (X\cos\theta - \frac{1}{\beta}Y\sin\theta),$$
$$m_{\beta+}(\theta) = \text{median } (\beta X\sin\theta + Y\cos\theta).$$

As before, it is straightforward to verify that the minimum attained by any angle $\theta \in [\pi/2, 2\pi]$ can be attained by an angle $\theta \in [0, \pi/2]$. Hence, it is without loss of generality to restrict attention to $\theta \in [0, \pi/2]$. Instead of $\theta = \pi/4$, the rotation that yields $m_{\beta-}(\theta) = 0$ is defined here by

$$\cos\theta = \frac{1}{\beta}\sin\theta \Leftrightarrow \theta = \arctan\beta.$$

We now show that Theorem 1 continues to hold:

Theorem 3 *Assume that X and Y are independent. The rotation with angle $\theta = 0$ is a local utility minimum if*

$$m_X f_X'(m_X) \geq 0, m_Y f_Y'(m_Y) \geq 0, \beta^2 m_X^2 + m_Y^2 \neq 0.$$

Proof. See Appendix A. ■

It is also straightforward to derive efficiency bounds. Here are two examples: 1) The universal bound based on the Hotelling-Solomons inequality (without any assumption on the underlying random variables) remains $\frac{1}{2}$, independently of β. 2) If X and Y are independent, using the generalized rotation where $\theta = \arctan\beta$, we obtain

$$EF(\arctan\beta) \geq \frac{1+\beta^2}{1+2\beta^2+(1-\beta^2)\cos^2(\arctan\beta)}.$$

We depict below the bound as a function of β (recall that $EF(\frac{\pi}{4}) \geq \frac{2}{3}$ with $\beta = 1$). Note that the bound tends back to the universal Hotelling-Solomons bound $\frac{1}{2}$ for $\beta \to 0$ and for $\beta \to \infty$. This is intuitive since in those limit cases one dimension becomes irrelevant and we obtain in the limit a one-dimensional voting problem where "rotations" cannot help.

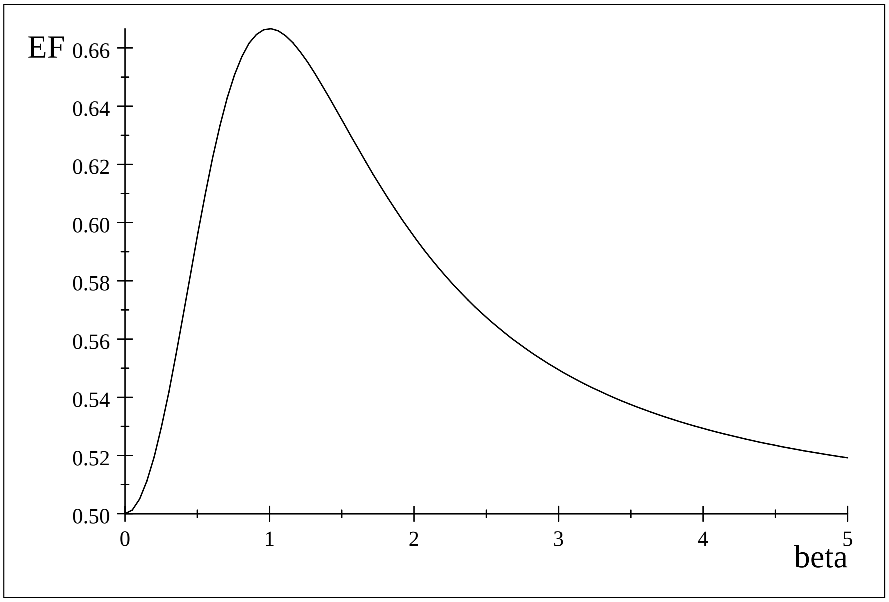

Figure 5. Bound on relative efficiency for I.I.D. random variables.

5.2 General Inner Product Norm

Consider next a general norm defined by an inner-product. Such a norm is generated by a symmetric, positive definite matrix Q:

$$||(x,y)|| \equiv \sqrt{(x,y)Q\ (x,y)^T}.$$

The "unit circle" is now an ellipse that is possibly tilted with respect to the standard coordinates. Let A_Q be the orthogonal matrix representing the change of variables that diagonalizes Q, and let M_Q be the obtained **diagonal** matrix.[33] Then M_Q is the matrix of a weighted Euclidean inner product, as explained above. The set of isometries that fix the origin and preserve orientation is thus given here by the composition:

$$A_Q R_{M_Q}\ (\theta) A_Q^{-1},$$

where $R_{M_Q}\ (\theta)$ is the set of generalized rotations that keep invariant the untilted unit ellipse associated to the diagonal matrix M_Q, as explained in the previous subsection. Note that the unit circle (i.e, ellipse) of this norm has now axes that are parallel to the coordinate axes defined by the change of variables A_Q. In particular, the relevant "zero rotation" is the one corresponding to these new variables; it is sub-optimal if the distribution of peaks has independent projections on these coordinates (rather than on the standard Cartesian ones).

[33] Note that any symmetric, positive definite matrix can indeed be diagonalized, and its two eigenvalues are always real.

6 Concluding Remarks

A re-definition of issues facilitates the search for consensus among ex-ante conflicting interests. We have shown that voting by simple majority on each dimension becomes a highly effective aggregation mechanism when combined with a judicious choice of the issues that are put up for vote. Our study endogenizes the process by which a "structure induced equilibrium" can be reached in a complex multi-dimensional collective decision problem with incomplete information about preferences. While we have focused on welfare maximization, other goals (such as maximizing the utility of an agenda setter) can be analyzed by the same methods.

7 Appendix A: Omitted Proofs

7.1 Proof of Theorem 1

In order to show that $\theta = 0$ is suboptimal, it is sufficient to show that

$$m_{-}(0)m'_{-}(0) + m_{+}(0)m'_{+}(0) = 0, \tag{11}$$

and that

$$m''_{-}(0)m_{-}(0) + (m'_{-}(0))^2 + m''_{+}(0)m_{+}(0) + (m'_{+}(0))^2 < 0. \tag{12}$$

By the definition of $m_{+}(\theta)$,

$$\begin{aligned} \frac{1}{2} &= F_{X\sin\theta + Y\cos\theta}(m_{+}(\theta)) \\ &= \int_{-\infty}^{\infty} \Pr\left(Y < \frac{m_{+}(\theta) - x\sin\theta}{\cos\theta}\right) f_X(x)\,dx \\ &= \int_{-\infty}^{\infty} F_Y\left(\frac{m_{+}(\theta) - x\sin\theta}{\cos\theta}\right) f_X(x)\,dx \end{aligned}$$

Since the above identity holds for all θ, we take the derivative with respect to θ and obtain

$$0 = \int_{-\infty}^{\infty} f_Y\left(\frac{m_{+}(\theta) - x\sin\theta}{\cos\theta}\right)\left(\frac{m'_{+}(\theta)\cos\theta - x + m_{+}(\theta)\sin\theta}{\cos^2\theta}\right) f_X(x)\,dx \tag{13}$$

Taking the second derivative with respect to θ, we obtain

$$\begin{aligned} 0 &= \int_{-\infty}^{\infty} f'_Y\left(\frac{m_{+}(\theta) - x\sin\theta}{\cos\theta}\right)\left(\frac{m'_{+}(\theta)\cos\theta - x + m_{+}(\theta)\sin\theta}{\cos^2\theta}\right)^2 f_X(x)\,dx \\ &+ \int_{-\infty}^{\infty} \frac{f_Y\left(\frac{m_{+}(\theta) - x\sin\theta}{\cos\theta}\right)}{\cos^4\theta}\left(\begin{array}{c}\left[m''_{+}(\theta)\cos\theta + m_{+}(\theta)\cos\theta\right]\cos^2\theta \\ +2\cos\theta\sin\theta\left(m'_{+}(\theta)\cos\theta - x + m_{+}(\theta)\sin\theta\right)\end{array}\right) f_X(x)\,dx \end{aligned} \tag{14}$$

If $\theta = 0$, then conditions (13) and (14) reduce to

$$0 = \int_{-\infty}^{\infty} f_Y(m_+(0)) \left(m'_+(0) - x\right) f_X(x)\, dx \tag{15}$$

and

$$0 = \int_{-\infty}^{\infty} f'_Y(m_+(0)) \left(m'_+(0) - x\right)^2 f_X(x)\, dx + \int_{-\infty}^{\infty} f_Y(m_+(0)) \left(m''_+(0) + m_+(0)\right) f_X(x)\, dx \tag{16}$$

Note that $m_+(0) = m_Y$, so it follows from (15) that

$$m'_+(0) = \frac{f_Y(m_Y) \int_{-\infty}^{\infty} x f_X(x)\, dx}{f_Y(m_Y) \int_{-\infty}^{\infty} f_X(x)\, dx} = \mu_X = 0,$$

and follows from (16) that

$$m''_+(0) = -m_Y - \frac{f'_Y(m_Y)}{f_Y(m_Y)} \int_{-\infty}^{\infty} x^2 f_X(x)\, dx.$$

Similarly, we can write

$$\frac{1}{2} = F_{X\cos\theta - Y\sin\theta}(m_-(\theta)) = \int_{-\infty}^{\infty} F_X\left(\frac{m_-(\theta) + y\sin\theta}{\cos\theta}\right) f_Y(y)\, dy$$

Taking the derivative with respect to θ, we obtain

$$0 = \int_{-\infty}^{\infty} f_X\left(\frac{m_-(\theta) + y\sin\theta}{\cos\theta}\right) \left(\frac{m'_-(\theta)\cos\theta + y + m_-(\theta)\sin\theta}{\cos^2\theta}\right) f_Y(y)\, dy$$

Taking the second derivative with respect to θ, we obtain

$$\begin{aligned} 0 = &\int_{-\infty}^{\infty} f'_X\left(\frac{m_-(\theta) + y\sin\theta}{\cos\theta}\right) \left(\frac{m'_-(\theta)\cos\theta + y + m_-(\theta)\sin\theta}{\cos^2\theta}\right)^2 f_Y(y)\, dy \\ &+ \int_{-\infty}^{\infty} \frac{f_X\left(\frac{m_-(\theta)+y\sin\theta}{\cos\theta}\right)}{\cos^4\theta} \left(\begin{array}{c} \left[m''_-(\theta)\cos\theta + m_-(\theta)\cos\theta\right]\cos^2\theta \\ +2\cos\theta\sin\theta\left[m'_-(\theta)\cos\theta + y + m_-(\theta)\sin\theta\right] \end{array}\right) f_Y(y)\, dy \end{aligned}$$

If $\theta = 0$, then the above two conditions reduce to

$$0 = \int_{-\infty}^{\infty} f_X(m_-(0)) \left(m'_-(0) + y\right) f_Y(y)\, dy$$

and

$$0 = \int_{-\infty}^{\infty} f'_X(m_-(0)) \left(m'_-(0) + y\right)^2 f_Y(y)\, dy + \int_{-\infty}^{\infty} f_X(m_-(0)) \left(m''_-(0) + m_-(0)\right) f_Y(y)\, dy$$

Since $m_-(0) = m_X$, we have

$$m'_-(0) = -\frac{\int_{-\infty}^{\infty} y f_Y(y)\, dy}{\int_{-\infty}^{\infty} f_Y(y)\, dy} = -\mu_Y = 0$$

and

$$m''_{-}(0) = -m_X - \frac{f'_X(m_X)}{f_X(m_X)} \int_{-\infty}^{\infty} y^2 f_Y(y)\, dy$$

Therefore, the first-order condition (11) holds because $m'_{-}(0) = m'_{+}(0) = 0$. For the second order condition (12), note that

$$\begin{aligned} & m''_{-}(0)m_{-}(0) + (m'_{-}(0))^2 + m''_{+}(0)m_{+}(0) + (m'_{+}(0))^2 \\ = \;& m_X \left(-m_X - \frac{f'_X(m_X)}{f_X(m_X)} \int_{-\infty}^{\infty} y^2 f_Y(y)\, dy\right) + m_Y \left(-m_Y - \frac{f'_Y(m_Y)}{f_Y(m_Y)} \int_{-\infty}^{\infty} x^2 f_X(x)\, dx\right) \\ = \;& -m_X^2 - m_Y^2 - m_X \frac{f'_X(m_X)}{f_X(m_X)} \int_{-\infty}^{\infty} y^2 f_Y(y)\, dy - m_Y \frac{f'_Y(m_Y)}{f_Y(m_Y)} \int_{-\infty}^{\infty} x^2 f_X(x)\, dx \end{aligned}$$

As a result, condition (12) is equivalent to

$$m_X^2 + m_Y^2 + m_X \frac{f'_X(m_X)}{f_X(m_X)} \int_{-\infty}^{\infty} y^2 f_Y(y)\, dy + m_Y \frac{f'_Y(m_Y)}{f_Y(m_Y)} \int_{-\infty}^{\infty} x^2 f_X(x)\, dx > 0.$$

Therefore, a sufficient condition for the sub-optimality of zero rotation is

$$m_X f'_X(m_X) \geq 0, m_Y f'_Y(m_Y) \geq 0 \text{ and } m_X^2 + m_Y^2 \neq 0.$$

7.2 Proof of Proposition 1

If X and Y are I.I.D., then we have

$$m_{-}(\pi/4) = 0 \text{ and } m_{+}(\pi/4) = \frac{\sqrt{2}}{2} m_{X+Y}.$$

Therefore, $\theta = \pi/4$ is a critical point if

$$0 = m_{-}(\pi/4)\, m'_{-}(\pi/4) + m_{+}(\pi/4)\, m'_{+}(\pi/4) = \frac{\sqrt{2}}{2} m_{X+Y} m'_{+}(\pi/4).$$

Recall (13) from the proof of Theorem 1 that

$$0 = \int_{-\infty}^{\infty} f_Y\left(\frac{m_{+}(\theta) - x\sin\theta}{\cos\theta}\right)\left(\frac{m'_{+}(\theta)\cos\theta - x + m_{+}(\theta)\sin\theta}{\cos^2\theta}\right) f_X(x)\, dx.$$

Hence, if X and Y are I.I.D. and $\theta = \pi/4$, we have

$$0 = \int_{-\infty}^{\infty} f_X\left(\sqrt{2}m_{+}(\pi/4) - x\right)\left(\sqrt{2}m'_{+}(\pi/4) - 2x + \sqrt{2}m_{+}(\pi/4)\right) f_X(x)\, dx.$$

It follows from the convolution of the distributions for X and Y that

$$\begin{aligned} & \sqrt{2}m'_{+}(\pi/4)\, f_{X+Y}\left(\sqrt{2}m_{+}(\pi/4)\right) \\ = \;& \int_{-\infty}^{\infty} f_X\left(\sqrt{2}m_{+}(\pi/4) - x\right)\left(2x - \sqrt{2}m_{+}(\pi/4)\right) f_X(x)\, dx. \end{aligned}$$

Note that by change of variable $y = \sqrt{2}m_+(\pi/4) - x$, we have

$$\begin{aligned}
&\int_{-\infty}^{\infty} f'_X\left(\sqrt{2}m_+(\pi/4) - x\right)\left(2x - \sqrt{2}m_+(\pi/4)\right)^2 f_X(x)\,dx \\
= &\int_{-\infty}^{\infty} f'_X(y)\left(2y - \sqrt{2}m_+(\pi/4)\right)^2 f_X\left(\sqrt{2}m_+(\pi/4) - y\right)dy.
\end{aligned}$$

Therefore,

$$\begin{aligned}
0 &= \int_{-\infty}^{\infty}\left[f'_X\left(\sqrt{2}m_+(\pi/4) - x\right)f_X(x) - f'_X(x)\,f_X\left(\sqrt{2}m_+(\pi/4) - x\right)\right]\left(2x - \sqrt{2}m_+(\pi/4)\right)^2 dx \\
&= \left[f_X\left(\sqrt{2}m_+(\pi/4) - x\right)f_X(x)\left(2x - \sqrt{2}m_+(\pi/4)\right)^2\right]_{-\infty}^{\infty} \\
&\quad - \int_{-\infty}^{\infty} f_X\left(\sqrt{2}m_+(\pi/4) - x\right)f_X(x)\,4\left(2x - \sqrt{2}m_+(\pi/4)\right)dx \\
&= 4\sqrt{2}m'_+(\pi/4)\,f_{X+Y}\left(\sqrt{2}m_+(\pi/4)\right)
\end{aligned}$$

where we use the assumption that

$$\begin{aligned}
&\lim_{x\to\infty} f_X\left(\sqrt{2}m_+(\pi/4) - x\right)f_X(x)\left(2x - \sqrt{2}m_+(\pi/4)\right)^2 \\
= &\lim_{x\to-\infty} f_X\left(\sqrt{2}m_+(\pi/4) - x\right)f_X(x)\left(2x - \sqrt{2}m_+(\pi/4)\right)^2 = 0.
\end{aligned}$$

Therefore, $m'_+(\pi/4) = 0$. It follows that $\sqrt{2}m_{X+Y}m'_+(\pi/4) = 0$, so $\theta = \pi/4$ is indeed a critical point.

7.3 Proof of Proposition 2

We prove the case where $m_X < \mu_X$. The case $m_X > \mu_X$ is analogous. In order to show that $\pi/4$ is globally optimal, it is sufficient to show for any $\beta \in [1/2, 1]$,

$$m_X \le m_{\sqrt{\beta}X_1 + \sqrt{1-\beta}X_2} \le m_{\frac{\sqrt{2}}{2}X_1 + \frac{\sqrt{2}}{2}X_2} < 0 = \mu_X. \tag{17}$$

Schur-concavity of $\Pr(X\sin\theta + Y\cos\theta \le z)$ in $(\sin^2\theta, \cos^2\theta)$ for all $\theta \in \left[0, \frac{\pi}{4}\right]$ and all $z \in [m_X, 0]$ is equivalent to

$$h(\beta) \equiv \Pr\left(\sqrt{\beta}X_1 + \sqrt{1-\beta}X_2 \le z\right) \text{ is weakly increasing in } \beta \text{ for all } z \in [m_X, 0]. \tag{18}$$

Now suppose condition (18) holds. It implies that

$$\Pr\left(\sqrt{\beta}X_1 + \sqrt{1-\beta}X_2 \le m_X\right) \le \Pr(X_1 \le m_X) = \frac{1}{2} \Rightarrow m_X \le m_{\sqrt{\beta}X_1 + \sqrt{1-\beta}X_2}.$$

Furthermore, (18) implies that, for all $\beta \in [1/2, 1]$,

$$\Pr\left(\frac{\sqrt{2}}{2}X_1 + \frac{\sqrt{2}}{2}X_2 \leq m_{\sqrt{\beta}X_1+\sqrt{1-\beta}X_2}\right) \leq \Pr\left(\sqrt{\beta}X_1 + \sqrt{1-\beta}X_2 \leq m_{\sqrt{\beta}X_1+\sqrt{1-\beta}X_2}\right) = \frac{1}{2}.$$

which implies

$$m_{\sqrt{\beta}X_1+\sqrt{1-\beta}X_2} \leq m_{\frac{\sqrt{2}}{2}X_1+\frac{\sqrt{2}}{2}X_2}.$$

Condition (17) then follows immediately, and thus the $\pi/4$-rotation is optimal.

7.4 Proof of Proposition 3

Suppose that $F_X(m_X + \varepsilon) + F_X(m_X - \varepsilon) \leq 1$ for all $\varepsilon > 0$. The other case is completely analogous. We first use an argument by van Zwet [1979] to claim that $m_X < \mu_X$. Note that

$$\begin{aligned} m_X - \mu_X &= \int_{-\infty}^{m_X} (m_X - x) f_X(x)\,dx + \int_{m_X}^{\infty} (m_X - x) f_X(x)\,dx \\ &= \int_{-\infty}^{m_X} F_X(x)\,dx - \int_{m_X}^{\infty} (1 - F_X(x))\,dx \\ &= \int_0^{\infty} [F_X(m_X - x) + F_X(m_X + x) - 1]\,dx \end{aligned}$$

It follows from $m_X \neq \mu_X$ that $m_X < \mu_X$. It also implies that $F_X(m_X - x) + F_X(m_X + x) - 1 < 0$ for some set of x with positive measure.

Next, we use an argument adapted from Watson and Gordon [1986] to prove that the median function is super-additive. The super-additivity of the median function is equivalent to

$$\Pr(X + Y < m_X + m_Y) < \frac{1}{2} \tag{19}$$

Note that

$$\begin{aligned} &\Pr(X + Y < m_X + m_Y) \\ =\ & \int_{m_Y}^{\infty}\int_{-\infty}^{m_X+m_Y-y} f_X(x) f_Y(y)\,dxdy + \int_{-\infty}^{m_Y}\int_{-\infty}^{m_X} f_X(x) f_Y(y)\,dxdy \\ & + \int_{m_X}^{\infty}\int_{-\infty}^{m_X+m_Y-x} f_X(x) f_Y(y)\,dydx \\ =\ & \int_{m_Y}^{\infty} F_X(m_X + m_Y - y) f_Y(y)\,dy + \frac{1}{4} + \int_{m_X}^{\infty} f_X(x) F_Y(m_X + m_Y - x)\,dx \\ =\ & \int_0^{\infty} F_X(m_X - \varepsilon) f_Y(m_Y + \varepsilon)\,d\varepsilon + \int_0^{\infty} f_X(m_X + \varepsilon) F_Y(m_Y - \varepsilon)\,d\varepsilon + \frac{1}{4} \end{aligned}$$

Therefore, condition (19) is equivalent to

$$4\int_0^\infty F_X(m_X-\varepsilon) f_Y(m_Y+\varepsilon) d\varepsilon + 4\int_0^\infty f_X(m_X+\varepsilon) F_Y(m_Y-\varepsilon) d\varepsilon < 1 \quad (20)$$

Let us define non-negative random variables X^+, X^-, Y^+, Y^- as

$$\begin{aligned} X^+ &= X - m_X | X \geq m_X \quad \text{and} \quad X^- = m_X - X | X \leq m_X \\ Y^+ &= Y - m_Y | Y \geq m_Y \quad \text{and} \quad Y^- = m_Y - Y | Y \leq m_Y \end{aligned}$$

Then

$$\begin{aligned} \Pr\left(X^- > Y^+\right) &= \int_0^\infty 2F_X(m_X-\varepsilon) 2f_Y(m_X+\varepsilon) d\varepsilon \\ \Pr\left(Y^- > X^+\right) &= \int_0^\infty 2F_Y(m_X-\varepsilon) 2f_X(m_X+\varepsilon) dx \end{aligned}$$

Therefore, condition (20) is equivalent to

$$\Pr\left(X^- > Y^+\right) + \Pr\left(Y^- > X^+\right) < 1 \quad (21)$$

A sufficient condition for (21) is

$$\Pr\left(X^+ < \varepsilon\right) \leq \Pr\left(X^- < \varepsilon\right) \text{ and } \Pr\left(Y^+ < \varepsilon\right) \leq \Pr\left(Y^- < \varepsilon\right) \quad (22)$$

for all $\varepsilon > 0$, and with strict inequality for some open interval of ε, because by setting $\varepsilon = Y^+$ and $\varepsilon = X^+$, respectively, we obtain

$$\Pr\left(X^+ < Y^+\right) < \Pr\left(X^- < Y^+\right) \text{ and } \Pr\left(Y^+ < X^+\right) < \Pr\left(Y^- < X^+\right)$$

and thus (21). Since X and Y are I.I.D., the sufficient condition (22) reduces to

$$\Pr\left(X^+ < \varepsilon\right) \leq \Pr\left(X^- < \varepsilon\right) \text{ for all } \varepsilon > 0.$$

Equivalently,

$$\Pr(X - m_X < \varepsilon) \leq \Pr(m_X - X < \varepsilon).$$

which simplifies into the first inequality in (9). As we argued above, since $m_X \neq \mu_X$, we must have $F_X(m_X-\varepsilon) + F_X(m_X+\varepsilon) - 1 < 0$ for some positive measure of ε, as desired.

Finally, we show that condition (10) ((9), respectively) is satisfied if F_X is strictly convex (concave, respectively). Note that $F(X)$ is uniformly distributed, so that $E[F(X)] = 1/2$. Suppose here that F is strictly convex. The concave case can be proved analogously. By Jensen's inequality

$$F(m_X) = \frac{1}{2} = E[F(X)] > F(E[X]) = F(\mu_X).$$

Hence, $m_X > \mu_X$. In order to show that $m_X + m_Y > m_{X+Y}$, it is sufficient to show that

$$F_X(m_X + \varepsilon) + F_X(m_X - \varepsilon) \geq 1 \text{ for all } \varepsilon > 0.$$

Note that $f_X(m_X + \varepsilon) - f_X(m_X - \varepsilon) > 0$ by strict convexity of F, so $F_X(m_X + \varepsilon) + F_X(m_X - \varepsilon)$ is increasing in ε and reaches a minimum at $\varepsilon = 0$. Since $F_X(m_X) + F_X(m_X) = 1$, we must have $F_X(m_X + \varepsilon) + F_X(m_X - \varepsilon) \geq 1$ for all $\varepsilon > 0$.

7.5 Examples for Section 3.3

We show here how the super-additivity condition of median is satisfied for two well-known families of distributions where condition (9) is not easily checked, or does not hold.[34]

Consider first the large and important family of gamma distributions with density

$$f_{\alpha,\beta}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \text{ for } x > 0.$$

This family contains the exponential (that can be obtained by setting $\alpha = 1$) and many other well known distributions. For any constant $c > 0$, the random variable cX is also gamma with parameters α and β/c. If X and Y are independent gamma with parameters (α_X, β) and (α_Y, β), respectively, then $X + Y$ is also gamma with parameters $(\alpha_X + \alpha_Y, \beta)$. Thus, the gamma family is closed under scaling and under convolution. In a classic study, Bock et al. [1987] showed that $\Pr(aX + bY \leq t)$, $0 \leq a, b \leq 1$, is Schur-convex in (a, b) for all $t \leq \mu_X$. Since $(1, 0) \succ \left(\frac{1}{2}, \frac{1}{2}\right)$, we have $F_{\frac{1}{2}X+\frac{1}{2}Y}(t) \leq F_X(t)$ for all $t \leq \mu_X$. Note that $m_X < \mu_X$ for gamma distributions (Groeneveld and Meeden [1977]), so we have $m_{\frac{1}{2}X+\frac{1}{2}Y} \geq m_X$ as desired.[35]

A second family is the Rayleigh distribution with cumulative distribution

$$F(x) = 1 - e^{-x^2} \text{ for } x \geq 0.$$

[34] Although the super-additivity (or sub-additivity) condition is derived for normalized distributions, it is straightward to verify that it is also sufficient for original distributions.

[35] Alternatively, let $m(\alpha, \beta)$ denote the median of gamma random variable X with parameters α and β. Then $m(\alpha, \beta) = m(\alpha, 1)/\beta$. Note that

$$\begin{aligned} U(\frac{\pi}{4}) &= -2\sigma^2(\alpha, \beta) - (\mu_+ - m_+)^2 \\ &= -2\sigma^2(\alpha, \beta) - \left(\frac{\sqrt{2}\alpha}{\beta} - \frac{\sqrt{2}}{2\beta} m(2\alpha, 1)\right)^2 \\ &= -2\sigma^2(\alpha, \beta) - \frac{1}{2\beta^2}(2\alpha - m(2\alpha, 1))^2 \end{aligned}$$

and

$$U(0) = -2\sigma^2(\alpha, \beta) - 2(\mu_X - m_X)^2 = -2\sigma^2(\alpha, \beta) - \frac{2}{\beta^2}(\alpha - m(\alpha, 1))^2.$$

Suppose X, Y are I.I.D. distributed according to Rayleigh.[36] Then, according to Lemma 4 in Hu and Lin [2000], we have

$$\Pr(X\cos\theta + Y\sin\theta \leq z) = 1 - \int_0^{\pi/2} \sin(2\tau)\left(1+\phi^2(\theta,\tau,z)\right)e^{-\phi^2(\theta,\tau,z)}d\tau$$

where $\phi(\theta,\tau,z) = z/\cos(\theta-\tau)$. The medians of X and of Y are $m_X = m_Y = \sqrt{\ln 2}$. It can be (numerically) verified that

$$\begin{aligned}\Pr\left((X+Y)/\sqrt{2} \leq \sqrt{2}m_X\right) &= 1 - \int_0^{\pi/2} \sin(2\tau)\left(1+\phi^2(\frac{\pi}{4},\tau,\sqrt{2\ln 2})\right)e^{-\phi^2(\frac{\pi}{4},\tau,\sqrt{2\ln 2})}d\tau \\ &\approx 0.4658 \\ &< 0.5 \\ &= \Pr\left((X+Y)/\sqrt{2} \leq m_+(\pi/4)\right)\end{aligned}$$

where the last equality follows from the definition of $m_+(\frac{\pi}{4})$. Hence, $m_+(\frac{\pi}{4}) > \sqrt{2}m_X$ as desired.

By assuming independence between X and Y, we were able to derive operational, sufficient conditions for the $\pi/4$ rotation to dominate the zero rotation, but independence is not necessary. We now present an example where, even though X and Y are correlated, the median function is super-additive (sub-additive) so the $\pi/4$ rotation is welfare superior to the zero rotation. The standard tool we use to model correlation between X and Y for given marginals is the copula (see Nelson [2006] for an introduction).

Example 3 *Suppose that X and Y are identically distributed on $[0,1]$ with marginals $F_X(x) = x^2$ and $F_Y(y) = y^2$. To model correlation between X and Y, we consider here the Farlie-Gumbel-Morgenstern (FGM) copula*

$$C_\delta(p,q) = pq + \delta pq(1-p)(1-q)$$

Therefore,

$$\begin{aligned}U(\frac{\pi}{4}) > U(0) &\Leftrightarrow \frac{1}{2\beta^2}(2\alpha - m(2\alpha,1))^2 < \frac{2}{\beta^2}(\alpha - m(\alpha,1))^2 \\ &\Leftrightarrow (2\alpha - m(2\alpha,1))^2 < 4(\alpha - m(\alpha,1))^2 \\ &\Leftrightarrow m^2(2\alpha,1) - 4\alpha m(2\alpha,1) < 4m^2(\alpha,1) - 8\alpha m(\alpha,1) \\ &\Leftrightarrow m(2\alpha,1) > 2m(\alpha,1).\end{aligned}$$

The last inequality holds because, as shown in Berg and Pedersen [2008], $m(\alpha,1)$ is convex in α.

[36] If Z_1, Z_2 is a random sample of size 2 from a normal distribution $N(0,1)$ then the distribution of $X = \sqrt{Z_1^2 + Z_2^2}$ is Rayleigh. In other words, the Rayleigh is the distribution of the norm of a two-dimensional random vector whose coordinates are normally distributed.

with $p, q \in [0,1]$ and $\delta \in [-1,1]$. The correlation coefficient for FGM copula is $\rho = \delta/3 \in [-1/3, 1/3]$. It follows from the Sklar theorem that we can write the joint distribution $F(x,y)$ in terms of its marginals and a copula $C(p,q)$:

$$F(x,y) = C(F_X(x), F_Y(y)).$$

With some algebra, we can derive the joint density as

$$f(x,y) = 4xy + 4\delta xy\left(2x^2-1\right)\left(2y^2-1\right).$$

Therefore, as in the proof of Proposition 3, we can write $\Pr(X+Y < m_X + m_Y)$ as

$$\begin{aligned}
& 2\int_{m_Y}^{1}\int_{0}^{m_X+m_Y-y} f(x,y)\,dxdy + \int_{0}^{m_Y}\int_{0}^{m_X} f(x,y)\,dxdy \\
= \; & 2\int_{\sqrt{2}/2}^{1}\int_{0}^{\sqrt{2}-y}\left(4xy + 4\delta xy\left(2x^2-1\right)\left(2y^2-1\right)\right)dxdy \\
& + \int_{0}^{\sqrt{2}/2}\int_{0}^{\sqrt{2}/2}\left(4xy + 4\delta xy\left(2x^2-1\right)\left(2y^2-1\right)\right)dxdy \\
= \; & \left(\frac{146}{35} - \frac{104}{35}\sqrt{2}\right)\delta - \frac{8}{3}\sqrt{2} + \frac{13}{3} \\
> \; & 0.5
\end{aligned}$$

for all $\delta \in [-1,1]$. Consequently, we have $m_{X+Y} < m_X + m_Y$. Since $F_X(x) = x^2$ is convex, $\mu_X < m_X$. It follows that the $\pi/4$-rotation dominates the zero-rotation in ex ante welfare. Alternatively, suppose $F_X(x) = \sqrt{x}$ and $F_Y(y) = \sqrt{y}$. If we again restrict attention to the FGM copula, we can follow the same procedure to show that $m_{X+Y} > m_X + m_Y$ and $\mu_X > m_X$.

7.6 Proof of Theorem 3

The proof follows the same steps as in proving Theorem 1. In order to show that $\theta = 0$ is suboptimal, it is sufficient to show

$$\beta^2 m_{\beta-}(0)m'_{\beta-}(0) + m_{\beta+}(0)m'_{\beta+}(0) = 0, \tag{23}$$

and

$$\beta^2 m''_{\beta-}(0)m_{\beta-}(0) + \beta^2(m'_{\beta-}(0))^2 + m''_{\beta+}(0)m_{\beta+}(0) + (m'_{\beta+}(0))^2 < 0. \tag{24}$$

By definition of $m_{\beta+}(\theta)$, we note that

$$\begin{aligned}
\frac{1}{2} &= F_{\beta X\sin\theta + Y\cos\theta}\left(m_{\beta+}(\theta)\right) \\
&= \int_{-\infty}^{\infty}\Pr\left(Y < \frac{m_{\beta+}(\theta) - \beta x\sin\theta}{\cos\theta}\right) f_X(x)\,dx \\
&= \int_{-\infty}^{\infty} F_Y\left(\frac{m_{\beta+}(\theta) - \beta x\sin\theta}{\cos\theta}\right) f_X(x)\,dx
\end{aligned}$$

Since it holds for all θ, we take a derivative with respect to θ to obtain

$$0=\int_{-\infty}^{\infty} f_Y\left(\frac{m_{\beta+}(\theta)-\beta x\sin\theta}{\cos\theta}\right)\left(\frac{m'_{\beta+}(\theta)\cos\theta-\beta x+m_{\beta+}(\theta)\sin\theta}{\cos^2\theta}\right)f_X(x)\,dx \tag{25}$$

By taking the second derivative with respect to θ, we obtain

$$\begin{aligned}0=&\int_{-\infty}^{\infty} f'_Y\left(\frac{m_{\beta+}(\theta)-\beta x\sin\theta}{\cos\theta}\right)\left(\frac{m'_{\beta+}(\theta)\cos\theta-\beta x+m_{\beta+}(\theta)\sin\theta}{\cos^2\theta}\right)^2 f_X(x)\,dx\\&+\int_{-\infty}^{\infty}\frac{f_Y\left(\frac{m_{+}(\theta)-\beta x\sin\theta}{\cos\theta}\right)}{\cos^4\theta}\left(\begin{array}{c}\left[m''_{\beta+}(\theta)\cos\theta+m_{\beta+}(\theta)\cos\theta\right]\cos^2\theta\\+2\cos\theta\sin\theta\left(m'_{\beta+}(\theta)\cos\theta-\beta x+m_{\beta+}(\theta)\sin\theta\right)\end{array}\right)f_X(x)\,dx\end{aligned} \tag{26}$$

If $\theta=0$, then conditions (25) and (26) reduce to

$$0=\int_{-\infty}^{\infty} f_Y\left(m_{\beta+}(0)\right)\left(m'_{\beta+}(0)-\beta x\right)f_X(x)\,dx$$

and

$$0=\int_{-\infty}^{\infty} f'_Y\left(m_{\beta+}(0)\right)\left(m'_{\beta+}(0)-\beta x\right)^2 f_X(x)\,dx+\int_{-\infty}^{\infty} f_Y\left(m_{\beta+}(0)\right)\left(m''_{\beta+}(0)+m_{\beta+}(0)\right)f_X(x)\,dx$$

Note that $m_{\beta+}(0)=m_Y$, so that we have

$$m'_{\beta+}(0)=\frac{\beta f_Y(m_Y)\int_{-\infty}^{\infty}xf_X(x)\,dx}{f_Y(m_Y)\int_{-\infty}^{\infty}f_X(x)\,dx}=\beta\mu_X=0,$$

and

$$m''_{\beta+}(0)=-m_Y-\frac{f'_Y(m_Y)}{f_Y(m_Y)}\int_{-\infty}^{\infty}\beta x^2 f_X(x)\,dx.$$

Similarly, we can write

$$\frac{1}{2}=F_{X\cos\theta-\frac{1}{\beta}Y\sin\theta}\left(m_{\beta-}(\theta)\right)=\int_{-\infty}^{\infty}F_X\left(\frac{m_{\beta-}(\theta)+\frac{1}{\beta}y\sin\theta}{\cos\theta}\right)f_Y(y)\,dy$$

Taking the derivative with respect to θ, we obtain

$$0=\int_{-\infty}^{\infty} f_X\left(\frac{m_{\beta-}(\theta)+\frac{1}{\beta}y\sin\theta}{\cos\theta}\right)\left(\frac{m'_{\beta-}(\theta)\cos\theta+\frac{1}{\beta}y+m_{\beta-}(\theta)\sin\theta}{\cos^2\theta}\right)f_Y(y)\,dy$$

Taking the second derivative with respect to θ, we obtain

$$\begin{aligned}0=&\int_{-\infty}^{\infty} f'_X\left(\frac{m_{\beta-}(\theta)+\frac{1}{\beta}y\sin\theta}{\cos\theta}\right)\left(\frac{m'_{\beta-}(\theta)\cos\theta+\frac{1}{\beta}y+m_{\beta-}(\theta)\sin\theta}{\cos^2\theta}\right)^2 f_Y(y)\,dy\\&+\int_{-\infty}^{\infty}\frac{f_X\left(\frac{m_{\beta-}(\theta)+\frac{1}{\beta}y\sin\theta}{\cos\theta}\right)}{\cos^4\theta}\left(\begin{array}{c}\left[m''_{\beta-}(\theta)\cos\theta+m_{\beta-}(\theta)\cos\theta\right]\cos^2\theta\\+2\cos\theta\sin\theta\left[m'_{\beta-}(\theta)\cos\theta+\frac{1}{\beta}y+m_{\beta-}(\theta)\sin\theta\right]\end{array}\right)f_Y(y)\,dy\end{aligned}$$

If $\theta = 0$, then the above two conditions reduce to

$$0 = \int_{-\infty}^{\infty} f_X\left(m_{\beta-}(0)\right)\left(m'_{\beta-}(0) + \frac{1}{\beta}y\right) f_Y(y)\,dy$$

and

$$0 = \int_{-\infty}^{\infty} f'_X\left(m_{\beta-}(0)\right)\left(m'_{\beta-}(0) + \frac{1}{\beta}y\right)^2 f_Y(y)\,dy + \int_{-\infty}^{\infty} f_X\left(m_{\beta-}(0)\right)\left(m''_{\beta-}(0) + m_{\beta-}(0)\right) f_Y(y)\,dy$$

Since $m_{\beta-}(0) = m_X$, we have

$$m'_{\beta-}(0) = -\frac{1}{\beta}\frac{\int_{-\infty}^{\infty} y f_Y(y)\,dy}{\int_{-\infty}^{\infty} f_Y(y)\,dy} = -\frac{1}{\beta}\mu_Y = 0$$

and

$$m''_{\beta-}(0) = -m_X - \frac{f'_X(m_X)}{f_X(m_X)}\int_{-\infty}^{\infty}\frac{1}{\beta^2}y^2 f_Y(y)\,dy$$

Therefore, the first-order condition (23) holds because $m'_{\beta-}(0) = m'_{\beta+}(0) = 0$. For the second-order condition (24), note that

$$\begin{aligned}
&\beta^2 m''_{\beta-}(0)m_{\beta-}(0) + \beta^2(m'_{\beta-}(0))^2 + m''_{\beta+}(0)m_{\beta+}(0) + (m'_{\beta+}(0))^2 \\
= \;& \beta^2 m_X\left(-m_X - \frac{f'_X(m_X)}{f_X(m_X)}\int_{-\infty}^{\infty}\frac{1}{\beta^2}y^2 f_Y(y)\,dy\right) + m_Y\left(-m_Y - \frac{f'_Y(m_Y)}{f_Y(m_Y)}\int_{-\infty}^{\infty}\beta^2 x^2 f_X(x)\,dx\right) \\
= \;& -\beta^2 m_X^2 - m_Y^2 - m_X\frac{f'_X(m_X)}{f_X(m_X)}\int_{-\infty}^{\infty} y^2 f_Y(y)\,dy - m_Y\frac{f'_Y(m_Y)}{f_Y(m_Y)}\int_{-\infty}^{\infty}\beta^2 x^2 f_X(x)\,dx
\end{aligned}$$

As a result, condition (24) is equivalent to

$$\beta^2 m_X^2 + m_Y^2 + m_X\frac{f'_X(m_X)}{f_X(m_X)}\int_{-\infty}^{\infty} y^2 f_Y(y)\,dy + m_Y\frac{f'_Y(m_Y)}{f_Y(m_Y)}\int_{-\infty}^{\infty}\beta^2 x^2 f_X(x)\,dx > 0.$$

Therefore, a sufficient condition for the sub-optimality of zero rotation is

$$m_X f'_X(m_X) \geq 0, m_Y f'_Y(m_Y) \geq 0 \quad \text{and} \quad \beta^2 m_X^2 + m_Y^2 \neq 0.$$

8 Appendix B: More than Two Dimensions

In this appendix, we sketch the generalizations of our main results (Theorems 1-2) to higher dimensions. Consider K independent issues, denoted by X_k, $k = 1, ..., K$. We write $\mathbf{X} = (X_1, ..., X_K)^T$ and assume that all random variables X_k are normalized. Let SO_K denote the *special orthogonal group* in dimension K which consists of $K \times K$ orthogonal matrices with determinant $+1$. This group is isomorphic to the set of rotations in $\mathbb{R}^K$. A $K \times K$ orthogonal matrix $Q \in SO_K$ is a real matrix with

$$Q^T Q = QQ^T = I$$

where Q^T is the transpose of Q, and where I is the $K \times K$ identity matrix. As a result

$$Q^{-1} = Q^T.$$

Each $K \times K$ special orthogonal matrix Q transforms an orthogonal system $\mathbf{X}$ into another orthogonal system while preserving the orientation in $\mathbb{R}^K$. The transformed orthogonal system $\mathbf{X}$ is denoted as $Q\mathbf{X}$. The planner's objective is to choose Q in order to maximize welfare.

8.1 The (Sub)-Optimality of the Zero- and $\pi/4$-Rotations

Theorem 1 can be easily extended to higher dimensions by applying our previous two-dimensional analysis to rotations of the first two dimensions only (while keeping all other dimensions fixed).

Suppose now that $X_1, ..., X_K$ are I.I.D. drawn from a common distribution. What is the counterpart of $\pi/4$ rotation (or equivalently the top-down procedure) in higher dimensions? We look for an orthogonal matrix Q that transforms $\mathbf{X}$ into a new vector $Q\mathbf{X}$ whose one coordinate is given by the sum $X_1 + ... + X_K$ while the other coordinates consists of various differences. For example, if $K = 4$, the orthogonal matrix Q (with determinant equal to $+1$) is given by

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} X_1 + X_2 - X_3 - X_4 \\ X_1 + X_4 - X_2 - X_3 \\ X_2 + X_4 - X_1 - X_3 \\ X_1 + X_2 + X_3 + X_4 \end{pmatrix}$$

More generally, consider an orthogonal matrix $\widehat{Q}$ with

$$\frac{1}{\sqrt{K}} \widehat{Q}_{ij} = \begin{cases} \text{either } 1 \text{ or } -1 & \text{for all } j \text{ if } \quad i \neq K \\ 1 & \text{for all } j \text{ if } \quad i = K \end{cases} \tag{27}$$

such that for all $k \neq K$, $\widehat{Q}_k\mathbf{X}$ contains an equal number of X_k's appearing with positive and negative signs. The matrix $\widehat{Q}_k$ is a *Hadamard matrix,* and the order of such a matrix must be $1, 2$ or a multiple of 4. Sylvester [1867] constructed Hadamard matrices of order 2^k for every non-negative integer k.[37] In those cases it is easy to see that the same condition we had before, namely the super-additivity of the median function, is again sufficient for the $\pi/4$-rotation to dominate the zero-rotation.

[37] The existence of Hadamard matrices of order $4k$ for every positive integer k is the well-known *Hadamard conjecture.* It was proven for all k up to 167.

8.2 Efficiency Bounds

As in the main text, we work here with the non-normalized random variables $X_1, ..., X_K$. With K dimensions, the expected utility of choosing marginal medians under an orthogonal transformation Q is given by

$$U(Q) = -\mathbb{E}\,||Q\mathbf{X} - \text{median}(Q\mathbf{X})||^2 = -\sum_{k=1}^{K} \text{var}(Q_k\mathbf{X}) - \sum_{k=1}^{K} (\text{mean}(Q_k\mathbf{X}) - \text{median}(Q_k\mathbf{X}))^2$$

where Q_k is the k-th row of the Q matrix. The first-best expected utility is $-\sum_{k=1}^{K}\text{var}(Q_k\mathbf{X})$. We define the relative efficiency of transformation Q relative to the first-best as:

$$EF(Q) \equiv \frac{\sum_{k=1}^{K} \text{var}(Q_k\mathbf{X})}{\sum_{k=1}^{K} \text{var}(Q_k\mathbf{X}) + \sum_{k=1}^{K} (\text{mean}(Q_k\mathbf{X}) - \text{median}(Q_k\mathbf{X}))^2}$$

Again, we can apply the Hotelling-Solomons inequality to obtain that

$$EF(I) \geq \frac{\sum_{k=1}^{K} \text{var}(Q_k\mathbf{X})}{\sum_{k=1}^{K} \text{var}(Q_k\mathbf{X}) + \sum_{k=1}^{K} \text{var}(Q_k\mathbf{X})} = \frac{1}{2}$$

Analogously, we can use the Basu-DasGupta inequality to show that, for unimodal distributions, we have

$$EF(I) \geq \frac{\sum_{k=1}^{K} \text{var}(Q_k\mathbf{X})}{\sum_{k=1}^{K} \text{var}(Q_k\mathbf{X}) + \frac{3}{5}\sum_{k=1}^{K} \text{var}(Q_k\mathbf{X})} = \frac{5}{8}$$

Now consider any even number K such that the Hadamard matrix exists. Suppose $X_1, ..., X_K$ are I.I.D. with log-concave densities. Consider again an orthogonal matrix $\widehat{Q}$ given in (27). It follows from the I.I.D. assumption that

$$\text{mean}(\widehat{Q}_k\mathbf{X}) - \text{median}(\widehat{Q}_k\mathbf{X}) = \begin{cases} 0 & \text{if } \ k \neq K, \\ \frac{1}{\sqrt{K}}(\text{mean}(\sum_{k=1}^{K} X_k) - \text{median}(\sum_{k=1}^{K} X_k)) & \text{if } \ k = K. \end{cases}$$

Therefore, we have

$$EF(\widehat{Q}) = \frac{\sum_{k=1}^{K} \text{var}(\widehat{Q}_k\mathbf{X})}{\sum_{k=1}^{K} \text{var}(\widehat{Q}_k\mathbf{X}) + \frac{1}{K}\left(\text{mean}(\sum_{k=1}^{K} X_k) - \text{median}(\sum_{k=1}^{K} X_k)\right)^2}$$

Given that $X_1, ..., X_K$ have log-concave densities, the convolution $Z \equiv \sum_{k=1}^{K} X_k$ also has a log-concave densities. Then the inequalities of Bobkov and Ledoux [2014] and of Ball and Böröczky [2010] together imply

$$(m_Z - \mu_Z)^2 \leq \frac{1}{f_Z^2(m_Z)} \ln^2\left(\sqrt{\frac{e}{2}}\right) \leq 12\sigma_Z^2 \ln^2\left(\sqrt{\frac{e}{2}}\right)$$

Hence,

$$EF(\widehat{Q}) \geq \frac{\sum_{k=1}^{K} \text{var}(\widehat{Q}_k \mathbf{X})}{\sum_{k=1}^{K} \text{var}(\widehat{Q}_k \mathbf{X}) + \frac{1}{K} 12\sigma_Z^2 \ln^2\left(\sqrt{\frac{e}{2}}\right)}$$

Let $\sigma_{X_k}^2$ denote the variance of X_k. Then we have $\sigma_Z^2 = K\sigma_{X_k}^2$ and

$$\text{var}(\widehat{Q}_k \mathbf{X}) = \widehat{Q}_k \widehat{Q}_k^T \sigma_{X_k}^2 = \sigma_{X_k}^2$$

since $\widehat{Q}_k \widehat{Q}_k^T = I$ by the definition of an orthogonal matrix. Therefore, we obtain the following efficiency bound for log-concave densities

$$EF \geq EF(\widehat{Q}) \geq \frac{K\sigma_{X_k}^2}{K\sigma_{X_k}^2 + 12\sigma_{X_k}^2 \ln^2\left(\sqrt{\frac{e}{2}}\right)} = \frac{1}{1 + \frac{1}{K} 12 \ln^2\left(\sqrt{\frac{e}{2}}\right)}$$

For example, if $K = 4$, the bound is 93.4%. Note that this bound is increasing the number of dimensions K, and tends to 100% when K goes to infinity.[38]

Remark 4 *More generally, consider any I.I.D. random variables $X_1, ..., X_K$ with finite means and variances and consider K such that the Hadamard matrix exists. Then, for the analog of the $\pi/4$ rotation we obtain that*

$$\begin{aligned} EF(\widehat{Q}) &= \frac{K\sigma_{X_k}^2}{K\sigma_{X_k}^2 + \frac{1}{K}(mean(\sum_{k=1}^{K} X_k) - median(\sum_{k=1}^{K} X_k))^2} \\ &= \frac{\sigma_{X_k}^2}{\sigma_{X_k}^2 + \frac{1}{K^2}(mean(\sum_{k=1}^{K} X_k) - median(\sum_{k=1}^{K} X_k))^2} \to 1 \text{ as } K \to \infty \end{aligned}$$

where the last assertion follows from the central limit theorem.

References

[2012] Ahn, D. and S. Oliveros (2012), "Combinatorial Voting," *Econometrica* **80**(1), 89–141.

[2005] Austen-Smith D. and J.S. Banks (2005), *Positive Political Theory II: Strategy and Structure*, The University of Michigan Press, Ann Arbor, MI.

[1966] Bahadur, R. R. (1966), "A Note on Quantiles in Large Samples," *Annals of Mathematical Statistics* **37**(3), 577-580.

[38] This asymptotic efficiency corresponds to the asymptotic efficiency of pure bundling when the number of objects goes to infinity and when marginal costs of production are zero (see Bakos and Brynjolfsson [1999]).

[1999] Bakos, Y. and E. Brynjolfsson (1999), "Bundling Information Goods: Pricing, Profits, and Efficiency," *Management Science* **45**(12), 1613-1630.

[2010] Ball, Keith M., and Karoly J. Boroczky (2010), "Stability of the Prekopa-Leindler Inequality," *Mathematika* **56**, 339-356.

[1993] Barbera, S., Gul, F. and Stacchetti, E. (1993), "Generalized Median Voter Schemes and Committees," *Journal of Economic Theory* **61**(2): 262-289.

[1965] Barlow, R. E., and F. Proschan (1965), *Mathematical Theory of Reliability,* J. Wiley & Sons. SIAM, Philadelphia, PA.

[1997] Basu, S. and A. DasGupta (1997), "The Mean, Median, and Mode of Unimodal Distributions: A Characterization," *Theory Probab. Appl.*, **41**(2), 210–223.

[2002] Battaglini, M. (2002), "Multiple Referrals and Multidimensional Cheap Talk," Econometrica 70(4), 1379-1401.

[2008] Berg, Christian, and Henrik L. Pedersen (2008), "Convexity of the Median in the gamma Distribution," *Ark. Mat.* **46**, 1–6.

[1948] Black, D. (1948), "On the Rationale of Group Decision-Making," *Journal of Political Economy* **56**(1), 23-34.

[1987] Bock M. E., P. Diaconis, F. W. Huffer and M. D. Perlman (1987), "Inequalities for Linear Combinations of gamma Random Variables," *Canadian Journal of Statistics* **15** (4), 387-395.

[2014] Bobkov, S. G., and Ledoux, M. (2014), "One-dimensional Empirical Measures, Order Statistics and Kantorovich Transport Distances," working paper, University of Minesota.

[1983] Border, Kim C., and J. S. Jordan (1983), "Straightforward Elections, Unanimity and Phantom Voters," *Review of Economic Studies* **50**(1), 153-170.

[1988] Caplin, Andrew, and Barry Nalebuff (1988), "On 64%-Majority Rule," *Econometrica* **56**(4), 787-814.

[1991] Caplin, Andrew, and Barry Nalebuff (1991), "Aggregation and Social Choice: A Mean Voter Theorem," *Econometrica* **59**(1), 1-23.

[2007] Charkraborty, A. and R. Harbaugh (2007), "Comparative Cheap Talk," *Journal of Economic Theory* 132, 70-94.

[1988] Dharmadhikari,S., and Kumar, J. (1988), *Unimodality, Convexity, and Applications*, Academic Press, First Edition.

[1988] Feld, S. and Grofman, B. (1988), "Majority Rule Outcomes and the Structure of Debate in One-Issue at a Time Decision-Making," *Public Choice* **59**, 239-252.

[1987] Ferejohn, J. and Krehbiel, K. (1987), "The Budget Process and the Size of the Budget," *American Journal of Political Science* **31**(2), 296-320.

[1907] Galton, F. (1907), "Vox Populi," *Nature* **75**, 450-451, 1949

[1997] Garcia-Roig, J. L. (1997): "On the group of isometries and a system of functional equations of a real normed plane." *Pitman Research Notes in Mathematics Series* **376**, 42-53.

[2017] Gershkov, A., Moldovanu, B. and Shi, X. (2017), "Optimal Voting Rules," *Review of Economic Studies*, **84**(2), 688-717.

[1977] Groeneveld, Richard A. and Glen Meeden (1977) "The Mode, Median, and Mean Inequality," *The American Statistician*, **31**(3), 120-121.

[1948] Haldane, J.B.S. (1948), "Note on the Median of a Multivariate Distribution," *Biometrika* **35**, 414-415.

[1932] Hotelling, Harold, and Leonard M. Solomons (1932), "The Limits of a Measure of Skewness," *Annals of Mathematical Statistics* **3**(2), 141-142.

[2000] Hu, Chin-Yuan, and Gwo Dong Lin (2000), "On an Inequality for the Rayleigh Distribution," *Sankhya: The Indian Journal of Statistics* **62**(1), 36-39.

[2001] Hu, Chin-Yuan, and Gwo Dong Lin (2001), "An Inequality for the Weighted Sums of Pairwise I.I.D. Generalized Rayleigh Random Variables," *Journal of Statistical Planning and Inference* **92**, 1-5.

[1956] Ibragimov, J. A. (1956). "On the Composition of Unimodal Distributions". *Teoriya Veroyatnostey* **I,** 283-288.

[2002] Kaas, R., Dhaene, J., Vyncke, D., Goovaerts, M.J., and Denuit, M. (2002), "A Simple Geometric Proof that Comonotonc Risks Have the Convex-Largest Sum," *Astin Bulletin* **32**(1), 71-80.

[1983] Karlin, S. and Rinott, Y. (1983). "Comparisons of measures, multivariate majorization and applications to statistics", in Studies in Econometrics, Time Series, and Multivariate Statistics (S. Karlin, T. Amemiya, L. A. Goodman, eds.). Academic Press, San Diego, 465-489.

[1984] Kim, K. H. and F. W. Roush (1984), “Nonmanipulation in Two Dimensions,” *Mathematical Social Sciences* **8**, 29-43.

[2017] Kleiner, A. and Moldovanu, B. (2017), “Content Based Agendas and Qualified-Majorities in Sequential Voting,” *American Economic Review* **107**(6), 1477-1506

[1972] Kramer, G. (1972), “Sophisticated Voting over Multidimensional Choice Spaces,” *Journal of Mathematical Sociology* **2**, 165-180.

[1973] Kramer, G. (1973), “On a Class of Equilibrium Conditions for Majority Rule,” *Econometrica* **41**(2), 285-297.

[1966] Lehmann, E. (1966), “Some Concepts of Dependence,” *Annals of Mathematical Statistics* **37**(5), 1137-1153.

[1980] Moulin, H. (1980), “On Strategy-Proofness and Single Peakedness,” *Public Choice* **35**, 437-455.

[1989] McAfee, P. ,J. McMillan and M. Whinston (1989): “Multiproduct Monopoly, Commodity Bundling, and Correlation of Values,” *Quarterly Journal of Economics* **104**(2), 371-383.

[2018] McMurray, Joseph C. (2018), “Why the Political World is Flat: An Endogenous ‘Left’ and ‘Right’ in Multidimensional Political Conflict,” working paper, Brigham Young University.

[2006] Neilsen, R.B. (2006), *An Introduction to Copulas*, second edition, Springer, New York.

[1983] Palfrey, T. (1983) “Bundling Decisions by a Multiproduct Monopolist with Incomplete Information,” *Econometrica* **51**(2), 463-483.

[1992] Peters, H., H. van der Stel, and T. Storcken (1992), “Pareto efficientity, Anonymity, and Strategy-Proofness in Location Problems,” *International Journal of Game Theory* **21**, 221-235.

[1993] Peters, Hans, Hans van der Stel, and Ton Storcken, "Generalized median solutions, strategy-proofness and strictly convex norms," *Zeitschrift für Operations Research* **38**.1 (1993): 19-53.

[1999] Poterba, J. M. and von Hagen, J. (eds), *Fiscal Institutions and Fiscal Perormance,* University of Chicago Press, Chiacgo.

[1973] Prekopa, A. (1973), "On Logarithmic Concave Measures and Functions," *Acta Sci. Math. (Szeged)* **34**, 335-343.

[1969] Rae, D. (1969), "Decision Rules and Individual Values in Constitutional Choice," *American Political Science Review* **63**, 40-56.

[1990] Small, Christopher G. (1990), "A Survey of Multidimensional Medians," *International Statistical Review,* **58**(3), 263-277.

[1867] Sylvester, J. (1867), "Thoughts on inverse orthogonal matrices, simultaneous sign successions, and tessellated pavements in two or more colours, with applications to Newton's rule, ornamental tile-work, and the theory of numbers," *Philosophical Magazine* 34, 461-475.

[1928] Szegö, G. (1928), "Über einige von S. Ramanujan gestellte Aufgaben," *Journal of the London Mathematical Society* **3**, 225-232.

[2000] Rychlik, T. (2000), "Sharp Mean-Variance Inequalities for Quantiles of Distribuition Determined by Convex and Star Orders," *Theory Probab. Appl.* **47**(2), 269-285.

[1979] Shepsle, K.A. (1979), "Institutional Arrangements and Equilibrium in Multidimensional Voting Models," *American Journal of Political Science* **23**, 27-60.

[1979] van Zwet, W.R. (1979), "Mean, Median, Mode II," *Statistica Neerlandica* **33** (1), 1-5.

[1986] Watson, R. and Gordon, I. (1986), "On Quantiles of Sums," *Australian Journal of Statistics* **28**(2), 192-199.